PHILOSOPHY OF MIND
PHILOSOPHY OF PSYCHOLOGY

PROCEEDINGS OF THE 9th INTERNATIONAL WITTGENSTEIN SYMPOSIUM
19th TO 26th AUGUST 1984, KIRCHBERG/WECHSEL (AUSTRIA)

PHILOSOPHIE DES GEISTES
PHILOSOPHIE DER PSYCHOLOGIE

AKTEN DES 9. INTERNATIONALEN WITTGENSTEIN SYMPOSIUMS
19. BIS 26. AUGUST 1984, KIRCHBERG/WECHSEL (ÖSTERREICH)

hpt

Digitized by the Internet Archive
in 2018 with funding from
Kahle/Austin Foundation

https://archive.org/details/philosophiedesge0000inte

PHILOSOPHIE DES GEISTES / PHILOSOPHIE DER PSYCHOLOGIE

PHILOSOPHY OF MIND / PHILOSOPHY OF PSYCHOLOGY

WITHDRAWN

hpt

SCHRIFTENREIHE DER WITTGENSTEIN-GESELLSCHAFT

Herausgeber / Board of editors

Elisabeth Leinfellner,
Rudolf Haller, Adolf Hübner, Werner Leinfellner, Paul Weingartner

VOLUME 11

BAND 11

PHILOSOPHY OF MIND
PHILOSOPHY OF PSYCHOLOGY

PROCEEDINGS OF
THE 9th INTERNATIONAL WITTGENSTEIN SYMPOSIUM

19th TO 26th AUGUST 1984
KIRCHBERG AM WECHSEL (AUSTRIA)

EDITORS

Roderick M. Chisholm, Johann Chr. Marek,
John T. Blackmore, Adolf Hübner

VIENNA 1985
HÖLDER-PICHLER-TEMPSKY

PHILOSOPHIE DES GEISTES
PHILOSOPHIE DER PSYCHOLOGIE

AKTEN DES
9. INTERNATIONALEN WITTGENSTEIN SYMPOSIUMS

19. BIS 26. AUGUST 1984
KIRCHBERG AM WECHSEL (ÖSTERREICH)

HERAUSGEBER

Roderick M. Chisholm, Johann Chr. Marek,
John T. Blackmore, Adolf Hübner

WIEN 1985
HÖLDER-PICHLER-TEMPSKY

Wir danken dem österreichischen Bundesministerium für Wissenschaft und Forschung und
der Kulturabteilung des Amtes der Niederösterreichischen Landesregierung
für die finanzielle Unterstützung bei der Drucklegung dieses Werkes.

BF
38
I55
484

Distributors for
Austria, Switzerland and the Federal Republic of Germany
VERLAG HÖLDER-PICHLER-TEMPSKY
A-1096 Wien, Postfach 127, Frankgasse 4

Distributors for the U.S.A. and Canada
KLUWER ACADEMIC PUBLISHERS GROUP
190 Old Derby Street, Hingham, MA 02043, U.S.A.

Distributors for all other countries
KLUWER ACADEMIC PUBLISHERS GROUP
Spuiboulevard 50, 3300 AZ Dordrecht, The Netherlands

CIP-Kurztitelaufnahme der Deutschen Bibliothek

Philosophie des Geistes, Philosophie der Psychologie:
Akten d. 9. Internat. Wittgenstein-Symposiums, 19.–26. August 1984, (Kirchberg/Wechsel, Österreich)/Hrsg. Roderick M. Chisholm . . . – Wien: Hölder-Pichler-Tempsky, 1985.
(Schriftenreihe der Wittgenstein-Gesellschaft; Bd. 11)
Parallelt.: Philosophy of mind, philosophy of psychology
ISBN 3-209-00592-3
NE: Chisholm, Roderick M. [Hrsg.]; Internationales Wittgenstein-Symposium <, 1984, Kirchberg, Wechsel >; Österreichische Ludwig-Wittgenstein-Gesellschaft: Schriftenreihe der Wittgenstein-Gesellschaft; PT

ISBN 3-209-00592-3

All Rights Reserved

Copyright © 1985 by HÖLDER-PICHLER-TEMPSKY, Vienna

No part of the material protected by this copyright notice may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording or by any informational storage and retrieval system without written permission from the copyright owner.

Gestaltung und Mitarbeit: Erich Péhm
Satz: Baroschrift, Wien
Druck: Manz, Wien

13173096

INHALTSVERZEICHNIS
TABLE OF CONTENTS

Vorwort/Preface 12
Abkürzungen der Titel von Wittgensteins Schriften /
Abbreviations of the Titles of Wittgenstein's Writings 21

RODERICK M. CHISHOLM: Eröffnungsrede 22
RODERICK M. CHISHOLM: Opening Address 25

1. Intentionalität
1. Intentionality

GUIDO KÜNG: Der intentionale und der reale Gegenstand 31
WERNER LEINFELLNER: Intentionality, Representation and the Brain Language 44
A. Z. BAR-ON: Who, Searle or Husserl? 56
RUDY KREJCI: Re-Emergence of the Concept of Consciousness in Twentieth Century Science and Philosophy 65
BURNHAM TERRELL: Intentionality and *Irrealia* 73
KARL WEICHSELBAUMER: Intentionalität in Technologie und teleologischer Erklärung . . 76
ELISABETH LEINFELLNER-RUPERTSBERGER: Intentions as *Intentiones Animae* in Ockham and as Immanent Mental 'Objects' in Brentano 79
AUSONIO MARRAS: Intentionality, Methodological Solipsism and Computational Psychology 86
ERIC RUSSERT KRAEMER: Function, Law and the Explanation of Intentionality 90
WOLFGANG KÜNNE: Die Intentionalität des Denkens 94
HJÖRDIS NERHEIM: Intentionalität bei Kant 98

2. Geist und Gehirn
2. Mind and Brain

DONALD M. MACKAY: The Costs of Mechanistic Embodiment 105
FRANZ SEITELBERGER: Gehirntätigkeit und Bewußtsein 112
PETER VITOUCH: Emotion und Sprache. Ein psychologischer Ansatz 115
FRANZ M. WUKETITS: Das Leib-Seele-Problem aus evolutionärer Perspektive 118
ANDREAS BARTELS: Zur Vereinbarkeit von Kripkes Theorie der notwendigen Identität mit der Leib-Seele-Identitätsthese 122
WOLFGANG WENNING: Vison, Brain, and Mind 126

3. Das Ich und die Welt
3. The Self and the World

JULIAN JAYNES: Four Hypotheses on the Origin of Mind 135
HERBERT STACHOWIAK: Über tautegorische Modelle 143
RODERICK M. CHISHOLM: A Logical Characterization of the Psychological 156
KONRAD CRAMER: Zeitlichkeit des Bewußtseins und synthetisches Urteil a priori 162
FRANCOIS BONSACK: Das Leib-Seele-Problem, epistemologisch gesehen 169
PAUL HÜBSCHER: Das metaphysische Subjekt: Die Stufe einer Leiter 175
ADOLF HÜBNER: On the Agreement of Human Mind and the Reason of Nature 178
GRAHAM P. CONROY: The Self, the World and the Metaphysical Subject 189
B. S. GOWER: The Self and Intentionality 192

4. Methodologie der Philosophie des Geistes
4. The Methodology of the Philosophy of Mind

HENRI LAUENER: Methodologische Bemerkungen zur Philosophie des Geistes 197
NICHOLAS RESCHER: Wie ist Naturwissenschaft möglich?
(Grundzüge eines naturalistischen Idealismus) 206
DIETER BIRNBACHER: Gibt es für das Leib-Seele-Problem eine „Lösung“? 215
ULRICH MAJER: Ramsey's Theory of Mind in Relation to his View on "Theories" 219
MICHAEL BENEDIKT: Leibniz' Representationsform eines adäquaten Begriffes
von ‚Geist' im Verhältnis zu Descartes und Kant 225
CHRISTFRIED TÖGEL: Zur Struktur psychologischer Theorien 231
EDWARD F. BECKER: Holistic Behaviorism 234
JOSEF MARSCHNER: Überlegungen zur Entstehung der Leib-Seele-Unterscheidung
und ihrem Zusammenhang mit Fragen der Erkenntnistheorie 237
NATIKA NEWTON: What Mental States are Really Like 240
TADEUSZ BUKSINSKI: Some Remarks on the Role of Indicators in Historical Inquiries . . . 243
MARK MIGOTTI: Metaphysics, Mind, and Physicalism 247
GARY FULLER: Instrumentalism in the Philosophy of Mind 251
SHARON LEE ARMSTRONG: Can Human Conceptual Schemes be Studied Empirically?
Some Cautionary Remarks 255
JOHN BLACKMORE: De-Idealizing the Mind-Body Problem 259
IZABELLA NOWAK & LESZEK NOWAK: The Position of the Theory of Mind
in the Categorial Interpretation of Dialectics 264
WILLIAM JAMES EARLE: Their Skulls and Our Thoughts 271
ROY BHASKAR: Weak Methods and Semantic Richness 274

5. Deskriptive Psychologie und Phänomenologie
5. Descriptive Psychology and Phenomenology

AARON BEN-ZEEV: The Description of Mental States 281
HANS KÖCHLER: Das Leib-Seele-Problem in der transzendentalen Phänomenologie 285
RAINER STUHLMANN-LAEISZ: The Uniqueness-Claim on Pleasure and its Empirical
Significance 291
FRANK W. LIEDKE: Intention und Erfüllung 294
PETER KLEIN: „Kultur“ und „Verhalten“ in der Philosophie des Geistes 297
BARRY SMITH: On the Structures of Perceptual Gestalten 301
WILLIAM BRENNER: Brown and Yellow 305
ROGER A. YOUNG: Knowing what it is like to be in a State of Mind 308
MANFRED SOMMER: Richard Avenarius und die Einheit von Weltbegriff
und Urempfindung 311
SARA YARETZKY-KAHANSKY: Descartes – Between Solipsism and Rationalism 315
KARL-ERNST BÜHLER: Ausdruck und psychosomatische Funktionen.
Eine empirische Untersuchung 319
JOHANN CHRISTIAN MAREK: Zur Analyse von ‚Träumen‘ 323

6. Handlung, Emotion und Wille
6. Action, Emotion and Will

BRIAN O'SHAUGHNESSY: The Varieties of Desire 333
KRYSTYNA GÓRNIAK-KOCIKOWSKA: Bewußtsein und der Wille zum Wollen 337
THOMAS KEUTNER: Absicht und Befehl 340
VEIKKO RANTALA: Intention and Action 343
REINER WIMMER: Aufweis des philosophischen Leib-Seele-Problems als eines Scheinproblems anhand des alltagssprachlich-lebensweltlichen Verständnisses der Leibgebundenheit unseres Handelns 348
ORA GRUENGARD: Unintended Actions and Unconscious Intentions 351
GUY A. M. WIDDERSHOVEN: Ego Development and Philosophy of Action 355
MICHAEL H. ROBINS: Intention and Control 358
ROBERT G. BURTON: Control 362

7. Intentionale Einstellungen
7. Intentional Attitudes

STEFAN SCHILLING: Zur Epistemologie intentionaler Einstellungen 367
PETER LUDLOW: Quantifying into Referential Attitudes 370
HAROLD MORICK: Sentences about Believing, II 373
STUART SILVERS: Explanation and Cognition 375
GERHARD HEYER: Bedingungen interner Repräsentation 378
DIETER STURMA: Ist Selbstbewußtsein eine Erkenntnis? 382

8. Wittgenstein und die Philosophie des Geistes
8. Wittgenstein and the Philosophy of Mind

GABRIEL FALKENBERG: Wittgenstein's Plan for the Treatment of Psychological Concepts . 387
MICHEL TER HARK: Mind as Expression 391
MASAHIRU OKU: In what Sense did Wittgenstein Reject the Mind-Body Problem? 394
J. C. NYÍRI: Wittgensteins Aufhebung der Gestalttheorie 399
H. E. MASON: The Notion of the Will in Wittgenstein's Later Writings 405
SYBE TERWEE: Malcolm's Dreaming and Wittgenstein's Philosophy of Mind 409
ANDREAS PUFF-TROJAN: ‚Ich bekenne, aus meinem Geist ein Idol gemacht zu haben, aber ich habe kein anderes gefunden.' (P. Valéry). – Bemerkungen zu Ludwig Wittgensteins Philosophie des Geistes und einem möglichen Parallelismus in der Gegenwartsliteratur . . . 414
WILHELM LÜTTERFELDS: Wittgensteins Rehabilitierung des Cartesianischen Paradigmas . 417
KENNETH JONES: The Tractatus Solution to the Problem of Consciousness 420
HANS RUDI FISCHER: Philosophie der Psychologie im *Tractatus* 423
PAUL HEDENGREN: An Illustration of Wittgenstein's Method 427

9. Wittgensteins frühe Philosophie
9. Wittgenstein's Early Philosophy

GILLES GASTON GRANGER: *Logisch-philosophische Abhandlung* et *Logischer Aufbau der Welt* 433
PETER M. SIMONS: Wittgenstein and the Semantics of Combination 446
GERT-JAN LOKHORST: A Formalization of Wittgenstein's Early Theory of Belief 450
MICHAEL V. WEDIN: Semantic Holism and the Phenomenalization of the *Tractatus* 453
ARTO SIITONEN: On the Concepts of Problem, Solution, Question and Answer in Wittgenstein's *Tractatus* 459
FU-TSENG LIU: Two Languages in Wittgenstein's *Tractatus*' Theory of Language 462
HEINZ HELLERER: Die Welt ist meine Sprache. Die idealistische Grundstruktur von Ludwig Wittgensteins *Logisch-philosophischer Abhandlung* 467
RÜDIGER E. BÖHLE: Ludwig Wittgenstein: *Tractatus* 6.52 oder das Dilemma wissenschaftlicher Aussagen 470

10. Wittgensteins späte Philosophie
10. Wittgenstein's Late Philosophy

ALDO GARGANI: Language and Community of Life in Wittgenstein's Later Philosophy . . . 477
FRIEDRICH WALLNER: Webern und Wittgenstein. Verbindlichkeit durch Elementarisierung . . . 482
SALMA SAAB: Sceptical Arguments and Psychological Concepts . . . 486
ARMIN BURKHARDT: Wittgensteins Wahrheit oder „Kann man die Wirkung der Erfahrung auf unser System von Annahmen leugnen?" (*ÜG* § 134) . . . 491
IRWIN GOLDSTEIN: Talking about Experiences . . . 496
STIG NYSTRAND: Private Experience and Public Communication . . . 499
ASTRID WINTERSBERGER: Wittgenstein – Ein Strukturalist . . . 505
BRUCE W. BROWER: Sensation Discourse and the Correctness Principle . . . 508
A. M. WILES: Theory and Phenomena: Plato and the Later Wittgenstein . . . 511
ELIZABETH H. WOLGAST: Certainty as a Form of Life . . . 515
NORBERT MEDER: Wittgenstein: Die Antizipation der Postmoderne . . . 519
MARGARETE MÖLLMANN: Wittgenstein on Mathematical Proof . . . 523
HORST RAMSENTHALER: Die Bedeutung Wittgensteins Spätphilosophie für das Verständnis familientherapeutischer Konzeptionen . . . 527
BENJAMIN F, ARMSTRONG, JR.: Wittgenstein on the Nature of Understanding . . . 531

11. Logik und die Philosophie der Sprache
11. Logic and the Philosophy of Language

HANS BURKHARDT: Modaltheorien Wittgensteins . . . 537
JERZY PERZANOWSKI: Some Observations on Modal Logics and the *Tractatus* . . . 544
ROLF GEORGE: Taking Thoughts one at a Time: Peirce on the Linear Progression of Thought in Reasoning . . . 551
GUGLIELMO TAMBURRINI: Turing, Gödel and Kalmár on Effective Procedures . . . 558
EDUARD W. WETTE: Inconsistent Polynominals, Geometrostatic and Theometric Locating of Monads, Tractatus-Residue . . . 562
PHILIP MAY: Mentally Represented Grammars and Computational Explanations of Understanding . . . 567
MARCELO DASCAL: Why should I Ask Her? . . . 571
AMADEO G. CONTE: Erkenntnis von Regeln und Erkenntnis durch Regeln . . . 577
KLAUS LANDWEHR: Wahrnehmung und Sprache: Evolutionäre Logik und historische Grammatik . . . 580
ULRICH STEINVORTH: How do Words Refer to Sensations? . . . 583
AUDUN ÖFSTI: Sprachspiel vs. vollständige Sprache . . . 586
HERBERT HRACHOVEC: Regeln und Intentionalität . . . 591
JOHN M. CONNOLLY: Frege, Sense, and Privacy . . . 594
RAINER P. BORN: „Realistische" Repräsentationen von „mentalen Phänomenen": Zur Situation der "situation semantics" . . . 597

12. Psychoanalyse
12. Psychoanalysis

H. STROTZKA: Psychoanalyse zwischen Psychologie und Philosophie . . . 603
WALTER TOMAN: Das Motivationskonzept der Psychoanalyse: Versuch einer wissenschaftstheoretischen Präzision mit Beispielen aus dem Alltag . . . 607
FRITZ J. KAUNE: Philosophische Fragen zur Archetypentheorie C. G. Jungs . . . 610
EUGENE T. GADOL: Category Mistakes of Freudianism and other False Ideologies (e.g. "Critical Rationalism" and Hermeneutics) . . . 615
WILJO DOELEMAN: I and My Neurosis: Psychoanalysis between Mechanism and Intentionality . . . 619

13. Universelle Sprachen
13. Universal Languages

IVO LAPENNA: Interaction between Language and Society 625
HELMAR G. FRANK: Experimente über Lernleichtigkeit und Informationsgehalt ILo-Vokabeln 630
LOTHAR HOFFMANN: Das semantische Mengengerüst für alle Sprachen 637
VLADIMIR MUŽIĆ: Pädagogische Relevanz einer universellen Sprache 640
JOHANN DVOŘAK: Universalsprache, Enzyklopädie und die „Allgemeine Verbesserung der menschlichen Dinge" bei J. A. Comenius 645
ECKEHART KÖHLER: On the Unity of all Normative Sciences as a Hierarchy of Rationality Principles 648
KARL MÜLLER: Zeitgemäße Wissenschaftsenzyklopädien 655
Liste und Index der Vortragenden/List and Index of the Speakers 658

VORWORT

Einleitung

Das Hauptthema des Neunten Internationalen Wittgenstein-Symposiums war „Philosophie des Geistes und Philosophie der Psychologie". U.a. wurden folgende Probleme im Detail diskutiert: das Problem von Geist und Körper (mind-body problem[1]), traditionelle Fragen der Philosophie des Geistes, die Philosophie der Psychologie und der „Funktionalismus".

Das Geist-Körper-Problem

Worin besteht das Problem, das wir im Deutschen gewöhnlich „Leib-Seele-Problem" und im Englischen „mind-body problem" nennen und das hier als „Geist-Körper-Problem" gekennzeichnet wird? Das englische Wort „mind" kann auf recht verschiedene Art und Weise verwendet werden, und diese Tatsache hat in den neueren philosophischen Diskussionen zu einiger Verwirrung geführt. Wir können mindestens fünf solche Gebrauchsweisen und daher mindestens fünf Gebrauchsweisen von „the mind-body problem" unterscheiden.

1. Wir könnten den Term „mind" so gebrauchen, wie Descartes den Term „mens" gebraucht hat, nämlich, daß er sich auf das bezieht, was psychologische Eigenschaften hat, auf das, was denkt, wahrnimmt, glaubt, wünscht. In diesem Falle würden wir „mind" so gebrauchen, daß der Term dasselbe wie „person" bedeutet, und daher so, daß er solche Entitäten wie den Verfasser und den Leser dieses Vorwortes bezeichnet. Wenn wir „mind" so gebrauchen, daß der Term dasselbe wie „person" bedeutet, dann würde eine Fassung des Geist-Körper-Problems wie folgende Frage aussehen: Was ist die Natur der Beziehung zwischen Personen und ihren Körpern? Was ist z.B. die Natur der Beziehung zwischen mir und meinem Körper? Es gibt zwei weitgespannte Möglichkeiten: Entweder bin ich mit meinem Körper identisch, oder ich bin mit meinem Körper nicht identisch. Wenn ich mit meinem Körper nicht identisch bin, dann gibt es wieder zwei Möglichkeiten: Entweder bin ich mit etwas, das einen Teil meines Körpers einschließt, identisch, oder ich bin nicht mit etwas, das einen Teil meines Körpers einschließt, identisch. Im letzteren Falle stellt sich die Frage: Was für ein Ding bin ich?

In den letzten Jahren ist vorgeschlagen worden, daß das Geistige, das, was im Englischen durch „mind" bezeichnet wird, mit dem Körper so verknüpft ist, wie das abstrakte Schema eines Computers mit dem Computer selbst. Aber dieser Vorschlag kann nicht richtig sein, wenn der Term „mind" in der eben herausgearbeiteten Bedeutung so aufgefaßt wird, daß er sich auf das bezieht, was psychologische Eigenschaften *hat*. Denn das abstrakte Schema eines Computers ist selbst eine Eigenschaft, ein abstraktes Objekt; jedoch das, was psychologische Eigenschaften hat, d.h. das, was wahrnimmt, denkt, fühlt und wünscht, ist ein individuelles Ding.

2. Wir könnten den Term „mind" auch so gebrauchen – wie es heute viele tun – daß er sich auf die intellektuellen Fähigkeiten einer Person bezieht. Das Wort wird auf diese Weise gebraucht, wenn man sagt, „That person has a good mind", und damit meint, daß die betreffende Person intelligent ist. Wenn wir diese zweite Bedeutung von „mind" im Auge haben, dann würde eine Untersuchung dessen, was durch „mind" bezeichnet wird, eine psychologische Untersuchung intellektueller Fähigkeiten sein. In diesem Falle könnte das Substantiv „mind" irreführend sein, denn es könnte so aufgefaßt werden, als bezeichnete es einen *individuellen Gegenstand*, der neben der Person und ihren Körper existiert. In solch einem Falle wird man dahingehend irregeführt, daß man annimmt, daß es *drei* individuelle Dinge gibt, die

zueinander in Beziehung gesetzt werden müssen: die Person, ihr Geistiges (mind) und ihr Körper.

Wenn wir „mind" auf diese zweite Art und Weise verwenden, können wir dann sagen, daß es ein Geist-Körper-Problem, ein mind-body-Problem gibt? Vielleicht taucht da die Frage auf: „Wie kann ein Körper intellektuelle Fähigkeiten haben?" Wenn wir annehmen können, daß es *etwas* gibt, das intellektuelle Fähigkeiten hat, dann kann diese Fassung des Geist-Körper-Problems auf die erste Fassung zurückgeführt werden: Worin besteht die Beziehung zwischen einer Person und ihrem Körper? Denn es ist anzunehmen, daß es die Person ist, die intellektuelle Fähigkeiten hat.

3. Wir können über den Geist einer Person auch so sprechen, daß er das ist, *womit* sie denkt. In diesem Fall bezeichnet der Term „mind" tatsächlich einen *individuellen Gegenstand*. Das, womit man denkt, ist ganz offensichtlich das Gehirn oder zumindestens etwas, das einen Teil des Gehirns einschließt. Wenn wir „mind" auf diese Art und Weise gebrauchen (d.h., wenn „mind" das bedeutet, womit man denkt), und nicht auf die erste Art und Weise (wo es die Person bedeutet), und auch nicht auf die zweite Art und Weise (wo es sich auf die intellektuellen Fähigkeiten bezieht), dann würde das, was wir eine „investigation of the mind", eine „Untersuchung des Geistes" nennen würden, eine neurophysiologische Untersuchung des Gehirns sein. Und der Ausdruck „mind-body problem" würde sich auf gewisse Fragen der Neurophysiologie beziehen. Die Antworten auf diese neurologischen Fragen liefern uns, für sich genommen, keine Lösung des Problems von Geist und Körper, des mind-body-Problems, wenn der Ausdruck „mind-body problem" im Einklang mit der *ersten* Bedeutung von „mind", die oben herausgearbeitet wurde, interpretiert wird.

4. Der Term „mind" wird manchmal dazu benützt, eine *geistige Substanz*, ein individuelles Ding von nicht-materiellem Charakter zu bezeichnen. (Das Wort „soul", „Seele", wäre hier vielleicht passender.) Wenn wir nun während der Beschäftigung mit dem Person-Körper-Problem die Entscheidung treffen würden, daß die Person nicht mit ihrem Körper identisch ist, und auch nicht mit etwas, das einen Teil ihres Körpers einschließt, dann könnten wir die Möglichkeit in Betracht ziehen, daß die Person eine Seele (soul) oder ein Geist, ein „mind" *ist*, in der gerade angeführten Bedeutung des Wortes „mind". (Wenn wir zu dem Ergebnis kommen, daß die Person eine Seele (soul) oder ein Geist (mind) *ist*, dann sollten wir vorsichtig sein und dieses Ergebnis nicht so ausdrücken, daß wir sagen, daß die Person eine Seele oder einen Geist *hat*. Denn dann könnten wir wieder dahingehend irregeführt werden, anzunehmen, daß wir es mit *drei* Arten von individuellen Dingen zu tun haben: mit Personen, ihrem Geist (minds) und ihren Körpern.)

5. Letztlich wird das Problem von Geist und Körper, das mind-body-Problem auch so aufgefaßt, daß es in der Untersuchung der Beziehung zwischen psychologischen und physischen Eigenschaften bestehe und daß dies ein besseres Verständnis nicht nur der Natur des psychologischen Universums, sondern auch der Natur des physischen Universums ermögliche. Man beschäftigt sich dann mit Fragen, wie: Ist jede psychologische Eigenschaft mit einer bestimmten physischen Eigenschaft *identisch*? Ist jede psychologische Eigenschaft in einem bestimmten Sinne von einer bestimmten physischen Eigenschaft oder von einer bestimmten Menge physischer Eigenschaften *abhängig*? Wiederum könnte hier der Ausdruck „*mind*-body problem" irreführend sein; ich schlage vor, daß „psychophysical problem", „psychophysisches Problem", besser wäre.

Die meisten gegenwärtigen Untersuchungen, die angeben, vom Geist-Körper-Problem zu handeln, befassen sich hauptsächlich mit dem psychophysischen Problem.

Deskriptive Psychologie

Ein großer Teil dessen, was „philosophy of mind" genannt wurde, befaßt sich damit, die psychologischen Zustände und Eigenschaften von Personen zu *beschreiben*, und zwar so zu beschreiben, wie sie von denen erlebt werden, die sich in diesen Zuständen befinden oder die

diese Eigenschaften haben. Ein Grund für dieses Vorgehen ist, daß man die philosophische Verwirrung, welche aus manchen dieser Zustände und Eigenschaften entspringt, auflösen will. Man hofft, dadurch die Person, die das Subjekt solcher Zustände und Eigenschaften ist, besser zu verstehen. Diese Art der Philosophie des Geistes schließt nicht nur Brentanos deskriptive Psychologie ein, sondern auch die „Phänomenologie" Husserls und der späteren Existentialisten, sowie auch diejenige Art der Analyse, die sich in Wittgensteins *Philosophischen Untersuchungen* findet.

Einige der Philosophen, die sich „Phänomenologen" genannt haben, haben gesagt, daß ihre Beschreibungen dafür bestimmt sind, das *Wesen* solcher Phänomene zu erfassen, zum Beispiel das Wesen eines solchen Phänomens wie des Traurigseins. Das Wesen eines *individuellen Dinges* würde dann eine Menge von Eigenschaften sein, die dieses Ding – und nur dieses Ding – notwendigerweise hat und die ein anderes Ding unter keinen Umständen haben könnte. Nun ist Traurigkeit eine *Eigenschaft* und nicht ein individuelles Ding. Sollen wir nun sagen, daß das Wesen der Eigenschaft Traurigkeit eine Eigenschaft der Eigenschaft Traurigkeit ist, eine Eigenschaft, die ihr – und nur ihr – wesentlich zukommt und die jedem anderen Ding völlig entgegengesetzt ist? Es würde dann nicht schwierig sein zu sagen, worin das Wesen der Traurigkeit besteht: Traurigkeit ist eine Eigenschaft, die notwendigerweise von der Art ist, daß alle Dinge – und nur die Dinge – die traurig sind, sie exemplifizieren.

Wenn Philosophen gesagt haben, daß sie das Wesen der Traurigkeit beschrieben haben, dann könnten sie damit gemeint haben, daß sie eine *Analyse* der Traurigkeit geliefert haben. Worin würde eine solche Analyse bestehen?

Vielleicht darin, daß sich die Philosophen mit einem wohlbekannten Phänomen befassen und daß sie gewisse wesentliche Eigenschaften dieses Phänomens aufzeigen, Eigenschaften, die einen tiefen Einblick gewähren, aber oft nicht bemerkt werden. Wenn jemand sagt: „Die Analyse der Traurigkeit muß ergeben, daß sie so-und-so beschaffen ist", und wenn ich verstehe, was gesagt wurde, dann bin ich entweder schon mit der Traurigkeit vertraut, oder fasse ich das, was der Sprecher sagt, so auf, daß das, mit dem *er* als Traurigkeit vertraut ist, so-und-so beschaffen ist. In beiden Fällen bin ich imstande, das Analysandum unabhängig zu identifizieren. Allgemeiner ausgedrückt: Wenn man angibt, daß man eine inhaltliche Aussage über die Natur eines Zustandes oder einer Eigenschaft macht, dann setzt man voraus, daß diese Eigenschaft unabhängig identifiziert werden kann. Aber manche Philosophen, die angeben, daß sie uns etwas über die „wahre Natur" des Psychologischen sagen, setzen dies nicht voraus.

Die Philosophen dieser Richtung sagen, daß die Daten, auf welche sich die deskriptiven Psychologen berufen, selbst zweifelhaft sind und daß sie daher mit den Altweibergeschichten einer laienhaften, volkstümlichen Psychologie verglichen werden können. Diese Philosophen machen sich offensichtlich Gedanken darüber, was psychologische Zustände und psychologische Eigenschaften sein würden, wenn 1. die Psychologie eine exakte Wissenschaft wäre, in dem Sinne, in dem man von der Physik und der Chemie sagen kann, daß sie exakte Wissenschaften sind, und wenn 2. die Psychologie trotzdem psychologische Zustände und Eigenschaften untersuchen würde. Man befaßt sich dann mit Fragen, wie: „Welche Arten von Dingen können gerechtfertigterweise ‚Glauben', ‚Fühlen', ‚Wahrnehmen' usw. genannt werden, wenn die Psychologie eine Naturwissenschaft wäre?" Manchmal wird vorgeschlagen, daß das, was dann engl. „psychological", dt. „psychologisch" oder engl. „mental", dt. „geistig" genannt werden könnte, von denjenigen psychologischen Zuständen und Eigenschaften, mit denen wir tatsächlich vertraut sind, sehr verschieden sein würde. In diesem Fall wäre es möglich, daß die Philosophen keine Theorie über die Natur derjenigen psychologischen Eigenschaften, mit denen *wir* unmittelbar vertraut sind, haben, denn sie könnten erklären, daß sie nicht an die Existenz solcher Eigenschaften glaubten. Es ist daher schwierig, ihre Aussage zu beurteilen, weil es nicht klar ist, was sie als ihre Daten ansehen oder was sie als Bestätigung oder als Falsifizierung dessen, was sie zu sagen versuchen, akzeptieren würden.

Die Begründer der „Wissenschaftlichen Weltauffassung“ und diejenigen, die ihnen in der Tradition des „Physikalismus“ folgten, hatten gehofft, imstande zu sein, psychologische Zustände und Eigenschaften mit physischen Zuständen und Eigenschaften zu identifizieren. Aber mit *welchen* physischen Zuständen und Eigenschaften? Für den Fall der Empfindungen schien das Programm recht vielversprechend zu sein, denn man dachte, daß hier jeder Typ von Empfindung ein eindeutiges physisches Korrelat hat. Man könnte dann als eine allgemeine Hypothese vorschlagen, daß jede Empfindung mit ihrem physischen Korrelat identisch ist. Aber es war schwierig, solche Korrelata für intentionale Phänomene (z.B. Glauben, Hoffen, Fragen, Wünschen) zu finden. Z.B. mit welchem physischen Zustand sollen wir den Wunsch, eine 23 Fuß lange Schaluppe zu besitzen, identifizieren? Es scheint so, als gäbe es keinen identifizierbaren physischen Zustand, der eindeutig mit einem solchen Wunsch korreliert ist.

Man kam daher auf den Gedanken, daß solche intentionale Phänomene, wenn schon nicht mit physischen *Zuständen* oder *Eigenschaften*, so doch mit *Verhaltensdispositionen* identifiziert werden könnten. Eine primitive Form dieser Ansicht wäre, daß man sich dann wünscht, eine 23 Fuß Schaluppe zu besitzen, wenn − und nur wenn − man auf die Frage, ob man eine 23 Fuß lange Schaluppe zu besitzen wünscht, bejahend antworten würde. Aber diese bikonditionale Aussage muß eingeschränkt werden. So könnte sie falsch sein, wenn sich jemand eine solche Schaluppe *nicht* wünscht, aber eine andere Person so irreführen möchte, daß sie glaubt, daß er sich eine solche Schaluppe wünscht. Und sie könnte falsch sein, wenn jemand sich *tatsächlich* solch eine Schaluppe *wünscht*, aber die Sprache dessen, der ihn fragt, nicht versteht. Ist es nun möglich, daß wir passende Einschränkungen finden und dieser bikonditionalen Aussage hinzufügen können? Die Schwierigkeit besteht darin, daß die Einschränkungen, die wir hinzufügen müssen, um solche Beispiele zu behandeln, sich auch auf *andere* intentionale Phänomene beziehen werden, z.B. auf andere Dinge, die man sich wünscht oder nicht wünscht, darauf, was man glaubt und was man nicht glaubt, darauf, was man wahrnimmt und woran man sich erinnert. Von solch einer Theorie kann man kaum sagen, daß sie psychologische Phänomene auf Verhaltensdispositionen reduziert, denn die relevanten Dispositionen können nur dann beschrieben werden, wenn man sich auf andere psychologische Phänomene bezieht.

Könnte man eine psychophysische Identitätsthese verteidigen, ohne sich auf eine psychologische *Korrelations*these zu berufen? Man kann natürlich nicht sagen, daß ein Gegenstand mit einem andern Gegenstand zu einer bestimmten Zeit identisch sein kann und dann zu einer anderen Zeit wiederum mit einem *anderen* Gegenstand. Aber es ist vorgeschlagen worden, daß die *Einzelvorkommen*, die „tokens“ einer beliebigen psychologischen Eigenschaft so beschaffen sein könnten, daß einige mit den Einzelvorkommen, den tokens, einer bestimmten physischen Eigenschaft identisch sind und andere mit den Einzelvorkommen, den tokens, einer anderen physischen Eigenschaft. Aber kann die metaphysische These, nach welcher es solche Einzelvorkommen, solche tokens, von Eigenschaften *gibt*, begründet werden? So weit ich sehen kann, hat die These überhaupt keine Begründung, ausgenommen die Möglichkeit, daß sie uns beim psychophysichen Problem helfen kann. Der metaphysische Preis dafür, daß man diese metaphysische Doktrin akzeptiert, um die psychophysische Identitätsthese zu retten, scheint sehr hoch zu sein, besonders wenn wir bedenken, daß diese „token-token“-Fassung der Theorie ein Paradebeispiel einer Theorie ist, die weder verifizierbar noch falsifizierbar ist.

Einige „Funktionalisten“ haben wirklich gesagt, daß eine Reduzierung des Psychologischen als Ganzes auf das Physische möglich ist, auch wenn eine Stück-für-Stück-Reduzierung unmöglich ist. Der springende Punkt ist dabei nicht, daß man spezifische psychologische Zustände (wie z.B., daß man sich eine 23 Fuß lange Schaluppe wünscht) mit neurologischen Zuständen oder mit Dispositionen, sich auf eine bestimmte, spezifische Art und Weise zu verhalten, identifiziert, sondern daß man seinen *gesamten* psychologischen Zustand in jeder

beliebigen Zeit mit seinem gleichzeitigen gesamten physischen Zustand identifiziert. Solch eine Theorie gehorcht nicht mehr dem Geist der „Wissenschaftlichen Weltauffassung", denn genau so wie die „token-token"-Theorie, die gerade erörtert wurde, scheint sie der paradigmatische Fall einer Theorie zu sein, die weder verifizierbar noch falsifizierbar ist.

* * *

Abschließend möchte ich im Namen der Herausgeber allen jenen danken, die an der Veröffentlichung dieses Bandes mitgewirkt haben: den Autoren der Beiträge; Frau Dr. Elisabeth Leinfellner, der Herausgeberin dieser Reihe; und Herrn Erich Péhm vom Verlag Hölder-Pichler-Tempsky. Wir danken auch dem Bundesministerium für Wissenschaft und Forschung, der Niederösterreichischen Landesregierung und der Österreichischen Ludwig Wittgenstein Gesellschaft für die gewährten Druckkostenzuschüsse, sowie Frau Helga Michelitsch für ihre umsichtige Hilfe bei der Vorbereitung für den Druck.

Roderick M. Chisholm

(ANMERKUNG DES ÜBERSETZERS)

[1] Während die traditionelle Übersetzung von „mind-body problem" „Leib-Seele-Problem" ist, wurden hier wegen der nur ungenauen semantischen Übereinstimmung der beiden Ausdrücke und des spezifischen Charakters dieses Vorworts die Übersetzung „Geist-Körper-Problem" für „mind-body problem" und „Geist" oder „Geistiges" für „mind" verwendet.

* * *

PREFACE

Introduction

The principal theme of the Ninth International Wittgenstein-Symposium is *Philosophy of Mind and Philosophy of Psychology*. Among the topics that are discussed in detail are: the mind-body problem; traditional questions of the philosophy of mind; and the philosophy of psychology.

The Mind-Body Problem

What is the mind-body problem? The word "mind" has a number of rather different uses and this fact has led to some confusion in recent discussions. We may distinguish at least five such uses and therefore at least five senses of "the mind-body problem."

(1) We could use the term "mind," as Descartes had used the term "*mens*," to refer to that which has psychological properties—to that which thinks, senses, believes, desires. In this case, we would be using "mind" to mean the same as "person" and hence to designate such entities as you and me. If we use "mind" to mean the same as "person," one form of the mind-body problem would be the question: What is the relation between persons and their bodies? What is the relation, for example, between me and my body? There are two broad possibilities: Either I am identical with my body or I am not identical with it. If I am not identical with my body, then once again there are two possibilities: either I am identical with something that includes a part of my body or I am not identical with anything that includes a part of my body. In the latter case, what kind of thing am I?

It has been suggested in recent years that "the mind" is related to the body in the way in which the abstract diagram of a computer is related to the computer. But this suggestion cannot be true if "mind" is taken, in the sense just distinguished, to refer to that which *has* psychological properties. For the abstract diagram of a computer is itself a property, an abstract object; but that which has psychological properties—that which senses, thinks, feels, and desires—is an individual thing.

(2) We could also use the term "mind," as many now do, to refer to a person's intellectual capacities. The word is being used in this way when one says, "That person has a good mind," meaning that the person is intelligent. An "investigation of the mind," in this second sense of the word "mind," would be a psychological investigation of intellectual capacities. In this case, the substantive "mind" may be misleading, since it might be taken to designate an *individual thing* that exists along with the person and his body. In such a case one is mislead into supposing that there are *three* individual things to be related—the person, the person's mind, and the person's body.

Using "mind" in this second way, may we say that there is a mind-body problem? Perhaps there is the question "How can a body have intellectual capacities?" If we may assume that there is *something* that has intellectual capacities, then this version of the mind-body problem reduces to the first: What is the relation between a person and his body? For it is presumably the person who has intellectual capacities.

(3) We may also speak of a person's mind as being *that by means* of which he thinks. In this case the term "mind" *does* designate an individual thing. That by means of which one thinks is quite obviously the brain—or at least something that includes a part of the brain. If we use "mind" in this way (to mean that by means of which one thinks) and not in the first way (to mean the person) and not in the second way (to refer to intellectual capacities), then what we

would call an "investigation of the mind" would be a neurophysiological investigation of the brain. And the expression "mind-body problem" would refer to certain questions of neurophysiology. Answers to these neurological questions do not, as such, give us a solution to the "mind-body problem," where this expression is interpreted in accordance with the *first* sense of "mind" distinguished above.

(4) The term "mind" is sometimes used to designate a *spiritual substance*, an individual thing of a nonmaterial nature. (Here, perhaps, the word "soul" is more appropriate.) If, in dealing with the person-body problem, we were to decide that the person is not identical with his body or with anything that includes a part of his body, then we might want to consider the possibility that the person *is* a soul or mind, in the present sense of the word "mind." (If we conclude that the person *is* a soul or mind, then we should take care not to express this conclusion by saying that the person *has* a soul or mind. For then we might be misled, once again, into thinking that we are dealing with *three* types of individual thing—persons, minds, and bodies.)

(5) Finally, "the mind-body problem" is sometimes taken to be that of studying the relation between psychological and physical properties and thereby getting a better understanding, not only of the nature of the psychological, but also that of the physical universe. The concern is with such question as: Is each psychological property *identical* with some physical property? Is each psychological property *dependent* in some special sense upon physical property or upon some set of physical properties? Here, once again, the expression "*mind*-body problem" may be misleading once again; I suggest that "psychophysical problem" would be better.

Most contemporary discussions that purport to be about the mind-body problem are concerned primarily with the psychophysical problem.

Descriptive Psychology

Much of what has been called "philosophy of mind" is concerned with *describing* the psychological states and properties of persons and describing them as they are experienced by those who are in those states or have those properties. One purpose of doing this is to resolve the philosophical perplexity to which some of these states and properties give rise. The hope is thereby better to understand the person who is the subject of such states and properties. This type of philosophy of mind includes, not only the descriptive psychology of Brentano, but also the "phenomenology" of Husserl and the later existentialists, as well as the type of study to be found in Wittgenstein's *Philosophical Investigations*.

Some of the philosophers who have called themselves "phenomenologists" have said that their descriptions are designed to catch the *essences* of such phenomena—the essence, say, of such a phenomenon as being sad. The essence of an *individual thing* would be a set of properties that that thing and only that thing has necessarily and that nothing else could possibly have. Now sadness is a *property* and not an individual thing. Shall we say, then, that the essence of the property sadness is a property of the property sadness—a property that is essential to it and only it and repugnant to every thing else? Then there would be no difficulty in saying what the essence of sadness is: sadness is a property which is necessarily such that it is exemplified in all and only those things that are sad.

When philosophers have said they were describing the essence of sadness, what they may have meant is that they were providing with an *analysis* of sadness. What would that be?

Perhaps this: the philosophers are concerned with a familiar phenomenon and that they are pointing out certain essential properties of it that are revealing but often go unnoticed. If you say, "Sadness is to be analysed as being so-and-so and such-and-such," and if I understand what you are saying, then either I am already acquainted with sadness or I take you to be saying that that which *you* are acquainted with as sadness is so-and-so and such-and-such. In either case, I am able to identify the *analysandum* independently. More generally, if one professes to give an informative statement about the nature of a state or property, then one pre-

supposes that that property can be independently identified. But some philosophers who profess to tell us about the "true nature" of the psychological do not presuppose this.

Such philosophers say that the data to which the descriptive psychologists appeal are themselves doubtful, being comparable to the old wives' tales of "folk psychology." These philosophers are apparently concerned with what psychological states and psychological properties would be if (1) psychology were a science in the sense in which physics and chemistry may be said to be science and if, nevertheless, (2) psychology were to investigate psychological states and properties. The concern, then, is with such questions as "What sorts of things might appropriately be called 'believing,' 'feeling,' 'sensing,' and the like if psychology were a natural science?" Sometimes it is suggested that what might then be called "psychological" or "mental" would be very different from the psychological states and properties that we are in fact acquainted with. In this case, the philosophers may have no theory about the nature of those psychological properties that *we* are directly acquainted with, since they may profess not to believe that there are such properties. It is difficult, therefore, to evaluate their statements since it is not clear what they take as their data or what they would accept as confirmation or as falsification of what they are trying to say.

The Psychophysical Problem

The authors of the *Wissenschaftliche Weltauffassung* and those who followed them in the tradition of "physicalism" hoped to be able to identify psychological states and properties with physical states and properties. But with *what* physical states and properties? The program seemed promising in the case of sensations where, it was thought, there is a physical correlate that is unique to every type of sensation. One could then propose as a general hypothesis that each sensation is identical with its physical correlate. But there was difficulty finding such correlates for intentional phenomena (e.g., believing, hoping, questioning, desiring). With what physical state, for example, are we to identify the desire to own a 23-foot sloop? There would seem to be no identifiable physical state that is uniquely correlated with such a desire.

It came to be thought, therefore, that such intentional phenomena might be identified, not with physical *states* or *properties*, but with *dispositions to behave*. A crude form of this view would be that you desire to own a 23-foot sloop if and only if you would respond affirmatively when asked whether you desire to own a 23-foot sloop. But this biconditional needs to be qualified. Thus it could be false if you do *not* desire such a sloop but wish to deceive someone into thinking that you do. And it could be false if you *do* desire such a sloop but do not understand the language of the one who is questioning you. Can we find suitable qualifications, then, to add to the biconditional? The difficulty is that the qualifications we must add to deal with such examples will refer to *other* intentional phenomena—for example, to other things that you desire or do not desire, to what you believe or do not believe, to what you perceive, and to what you remember. Such a theory could hardly be said to reduce psychological phenomena to dispositions to behave—since the relevant dispositions can be described only by reference to other psychological phenomena.

Could one defend a psychophysical identity thesis without appeal to a psychophysical *correlation* thesis? One cannot say, of course, that a thing may be identical with one thing at one time and with *another* thing at another time. But it has been suggested that the *particular occurrences* ("tokens") of any given psychological property may be such that some are identical with the particular occurrences of one physical property and others are identical with the particular occurrences of some other physical property. But what ground is there for the metaphysical thesis according to which there *are* such particular occurrences (tokens) of properties? So far as I have been able to see, the thesis has no basis at all—other than the possibility that it might help us with the psychophysical problem. The metaphysical price of accepting this metaphysical doctrine in order to preserve the psychophysical identity theory seems very

high, especially when we consider that this "token-token" version of the theory is *par excellence*, a theory which is neither verifiable nor falsifiable.

Some "functionalists" have said, in effect, that even though a piecemeal reduction of the psychological to the physical is not possible, a wholesale reduction is possible. The point is, not to identify particular psychological states (such as desiring a 23-foot sloop) with neurological states or with dispositions to behave in certain particular ways, but to identify one's *total* psychological state at any time one's total physical state at that time. Such a theory is no longer in the spirit of the *Wissenschaftliche Weltauffassung*, for, like the "token-token" theory just considered, it would seem to be a paradigm case of a theory that is neither verifiable nor falsifiable.

* * *

On behalf of the Board of Editors, I would like to thank all of those who have worked with us in preparing this volume for publication: the authors of the particular papers; Dr. Elisabeth Leinfellner, one of the General Editors of this series; and Mr. Erich Péhm of Hölder-Pichler-Tempsky. We also thank the Bundesministerium für Wissenschaft und Forschung, the Niederösterreichische Landesregierung and the Austrian Ludwig Wittgenstein Society for grants to cover the cost of publication and Mrs. Helga Michelitsch for her invaluable help in preparing the final manuscript.

Roderick M. Chisholm

* * *

ABKÜRZUNGEN DER TITEL VON WITTGENSTEINS SCHRIFTEN
ABBREVIATIONS OF THE TITLES OF WITTGENSTEIN'S WRITINGS

BBB = *Blue and Brown Books*

BF = *Bemerkungen über Farben*

BLB = *Blue Book*

BPP = *Bemerkungen über die Philosophie der Psychologie*

BRB = *Brown Book*

BFGB = *Bemerkungen über Frazers „The Golden Bough"*

BGM = *Bemerkungen über die Grundlagen der Mathematik*

CV = *Culture and Value*

LCA = *Lectures and Conversations on Aesthetics, Psychology and Religious Belief*

LE = *Lecture on Ethics*

LFM = *Lectures on the Foundations of Mathematics*

LO = *Letters to Ogden*

LRKM = *Letters to Russell, Keynes and Moore*

LS = *Letzte Schriften über die Philosophie der Psychologie*

LW = *Last Writings on the Philosophy of Psychology*

NB = *Notebooks 1914–1916*

NL = *Notes for Lectures on Private Experience and Sensedata*

OC = *On Certainty*

PB = *Philosophische Bemerkungen*

PG = *Philosophische Grammatik / Philosophical Grammar*

PI = *Philosophical Investigations*

PR = *Philosophical Remarks*

PT = *Prototractatus*

PU = *Philosophische Untersuchungen*

RC = *Remarks on Color*

RFM = *Remarks on the Foundations of Mathematics*

RLF = *Some Remarks on Logical Form*

RPP = *Remarks on the Philosophy of Psychology*

TLP = *Tractatus Logico-Philosophicus*

TB = *Tagebücher 1914–1916*

ÜG = *Über Gewißheit*

VB = *Vermischte Bemerkungen*

WV = *Wörterbuch für Volksschulen*

WWK = *Wittgenstein und der Wiener Kreis*

Z = *Zettel*

ERÖFFNUNGSREDE

Roderick M. Chisholm
Brown University, Providence, R.I.

Ich danke den Mitgliedern des Programmkomitees, daß sie mir die Ehre gegeben haben, diese Eröffnungsbemerkungen zu machen.

Es gibt viele verschiedene Weisen, wie wir die Arbeit eines großen Philosophen betrachten können. Um den Ausdruck von Wittgenstein zu gebrauchen: wir können sie unter sehr verschiedenen „Aspekten" verstehen. Ich will heute morgen nur von einem jener Aspekte sprechen, unter denen wir das Werk von Wittgenstein betrachten können. Diese Perspektive ist historisch bedeutend; ich glaube, sie wird seine Philosophie beleuchten; und, soviel ich weiß, hat niemand sie ausführlich diskutiert. Sie besteht darin, daß wir das Werk Wittgensteins als eine Fortsetzung von dem von Franz Brentano und seiner Schule betrachten.

Brentano hat einmal eine Abhandlung über Thomas Reid geschrieben. Er hat sie in zwei Teilen formuliert – „was an Reid zu tadeln ist" und „was an Reid zu loben ist". Würde er über Wittgenstein schreiben, so könnten wir es nicht verhindern, daß er einen langen Teil schriebe über das, was an Wittgenstein zu tadeln ist. Aber es würde nicht schwierig sein, ihn zu überreden, einen längeren Teil über das, was an Wittgenstein zu loben ist, zu schreiben.

Um die Perspektive, von der ich jetzt spreche, zu verstehen, sollte man mit Wittgensteins *Bemerkungen über Farben* anfangen. Sein Büchlein ist, ebenso wie Brentanos *Untersuchungen zur Sinnespsychologie*, eine Abhandlung über die Gegenstände der Empfindung. Die Forschung auf diesem Gebiet war ein gemeinsamer Zug der Brentano-Schule und des österreichischen Positivismus – insbesondere von Mach, Schlick und dem frühen Carnap. Sie wurde von den Positivisten aufgegeben, als sie sich dem Physikalismus zuwandten, aber sie wurde nie von Wittgenstein aufgegeben. Was Wittgenstein und Brentano über Farben sagen, wäre inakzeptabel für einen Philosophen, der bekannte, er wisse nicht was eine Empfindung sei. Die Fragen, die sie besprechen, haben überhaupt nichts mit Physik oder mit physikalischen Körpern zu tun. Die Antwort auf eine Frage dieser Art kann, wie Wittgenstein sich ausdrückt, „kein physikalistischer Satz sein."

Sein Werk über Farben, sagte Wittgenstein, ist nicht „eine *Theorie* der Farben (weder eine physiologische noch eine psychologische), sondern die *Logik des Farbbegriffes*" (S. 15). Sicherlich hat Wittgensteins Werk größtenteils mit Begriffen zu tun. Wenn er bekennt, über *Sprache* zu sprechen, spricht er oft nicht über Sprache, sondern über Begriffe. Tatsächlich hätte er, wo er den Ausdruck „Sprachspiel" verwendet, mit gleicher Berechtigung den Ausdruck „*Begriffs*spiel" verwenden können.

Wittgenstein sagt in den *Bemerkungen über Farben*: „Hier ist die Versuchung sehr groß, an eine Phänomenologie, ein Mittelding zwischen Wissenschaft und Logik, zu glauben" (S. 35). Ist er selbst dieser Versuchung erlegen? Er sagt: „Es gibt zwar keine Phänomenologie, wohl aber phänomenologische Probleme" (S. 22). Und dies sind die Probleme, die er diskutiert.

Die *Philosophischen Untersuchungen* sind wohl ein bahnbrechender Beitrag zu dem, was Brentano „deskriptive Psychologie" genannt hatte. Sie erweitern unser Verständnis der Intentionalität in einem Maß, das die Mitglieder der Brentano-Schule nie hätten leisten können. Um diesen Punkt einzusehen, braucht man nur zu betrachten, was Wittgenstein uns dort über intentionale Phänomene erzählt – insbesondere über Wissen, Glauben, Wahrnehmen und Wollen. Und wenn er von dem „Aufleuchten eines Aspektes" spricht, macht er einen tiefschürfenden Beitrag zu unserem Verständnis von *Bemerken* – einem Begriff, der eine grundlegende Rolle in Brentanos deskriptiver Psychologie spielt.

Dieses Aufleuchten eines Aspektes ist etwas, was wir hervorheben müssen, denn es gibt eine auffallende Ähnlichkeit zwischen dem, was Wittgenstein hier geleistet hat, und dem, was Brentano geleistet hatte, als er unser Bewußtsein von zeitlichen Verhältnissen besprochen hatte.

Andere Philosophen hatten bemerkt, daß unser Zeitbewußtsein philosophisch rätselhaft ist. Man kann also fragen: „Wenn du ein Ding als zeitlich *vor* einem anderen stehend wahrnimmst, so mußt du das erste Ding als vergangen wahrnehmen. Aber wenn ein Ding jetzt in der Vergangenheit ist und nicht mehr in der Gegenwart, wie kannst du es jetzt *wahrnehmen* – und es *als vergangen* wahrnehmen?“ Meines Erachtens war es Brentano, der als erster die volle Bedeutung solcher Fragen gesehen hatte. Er sah, daß sie eine Umwandlung fundamentaler psychologischer und philosophischer Begriffe verlangen. Das Ergebnis war eine Theorie des Zeitbewußtseins – ein Beitrag, welcher, zum Glück oder Unglück, zu dem Werk Heideggers und seiner Nachfolger geführt hat.

Ähnlich ist es mit Wittgenstein und dem Aufleuchten eines Aspektes. Psychologen wußten, daß das Beispiel des Hasen-Enten-Bildes uns vor rätselhafte Fragen stellt. Etwa: „Wenn du es als Ente zu sehen aufhörst und es als Hasen zu sehen anfängst, was ist es, das sich ändert – der *Gegenstand* der Empfindung oder der *Akt* der Empfindung?“ Es war Wittgenstein, der als erster die volle Bedeutung solcher Fragen sah. Und was er darüber sagt, ist, meiner Meinung nach, noch nicht völlig gewürdigt worden. Von dem in Frage kommenden Phänomen sagt er: „Seine *Ursachen* interessieren den Psychologen, nicht mich“ (*LS* 434). Auch hier stellt er Fragen deskriptiver Psychologie zur Diskussion.

Es gibt natürlich wichtige Unterschiede zwischen Brentano und Wittgenstein. Wir dürfen sie nicht unterschätzen. Wir sollten sie aber auch nicht überschätzen.

Ich denke zum Beispiel an die Erkenntnistheorie – an die Evidenzlehre, wie man sie in der Brentano-Schule genannt hat. Was Wittgenstein über Evidenz sagt – insbesondere in *Über Gewißheit* – läßt sich eher als eine Ergänzung denn als ein Widerspruch zur Evidenzlehre Brentanos interpretieren.

Wittgenstein sagt vieles über unmittelbare Erfahrung, was Brentano früher behauptet hatte. Zum Beispiel: „Man kann sich freilich über die unmittelbare Erfahrung nicht täuschen.“ Und dann, anders als Brentano, fügt er hinzu: „aber nicht, weil sie so gewiß ist.“ (*LS* 187) Wo Brentano sagt „Ich weiß“ und „Es ist evident“, sagt Wittgenstein: „Es hat keinen Sinn zu sagen ‚Ich weiß‘ oder ‚Es ist evident.‘“ Aber man kann sich des Eindrucks nicht erwehren, daß der Streit nur sprachlich ist. Die Kriterien, durch welche Wittgenstein entscheidet, daß „Ich weiß“ und „Es ist evident“ *sinnlos* sind, sind genau dieselben wie die Kriterien, durch welche Brentano entscheidet, daß sie *wahr* sind.

Wenn eine Person Schmerzen hat, *sagt* sie wahrscheinlich nicht „Es ist mir evident, daß ich Schmerzen habe“. Aber sie ist imstande, die Tatsache, daß sie Schmerzen hat, *als Evidenz* zu gebrauchen – als Evidenz für eine *andere* Tatsache. Um diesen Punkt besser zu verstehen, betrachten wir den folgenden Dialog:

„Offensichtlich habe ich mich noch nicht erholt.“

„Woher weißt du das?“

„Ich habe noch Schmerzen.“

Unabhängig davon, ob wir von einer Person sagen, daß sie *wisse, daß sie Schmerzen hat, können wir sagen, sie verwendet* diese Tatsache *als Evidenz* für die Hypothese, daß sie sich noch nicht erholt hat. Es scheint so, wie wenn der eine Philosoph sagt, „Es gibt einen unbewegten ersten Beweger,“ und der andere sagt, „Es gibt einen ersten Beweger, der sich selbst bewegt.“ Und doch sind die Gründe in beiden Fällen dieselben.

Es ist wahr, daß dort, wo Brentano das *Denken* betont, Wittgenstein *das Sprechen* hervorhebt. Brentano legt es uns nahe, daß die Prinzipien der Logik Regeln des *Denkens* sind. Und Wittgenstein legt nahe, daß sie Regeln der *Sprache* sind. Wenn man den einen Philosophen des “Psychologismus“ bezichtigen kann, so kann man den anderen des „Lingualismus“ bezichtigen. Indes scheinen mir die Anklagen in beiden Fällen nicht gerechtfertigt zu sein. Wie entscheidet man, *was* die Regeln des Denkens oder was die Regeln der Sprache sind?

Was der eine Philosoph sagen kann, kann der andere Philosoph auch sagen – *mutatis mutandis*.

Sollten wir die Sprache in bezug auf das Denken verstehen – oder umgekehrt? Die Antwort dreht sich um unsere Position zu einer der fundamentalsten aller Fragen, die in der österreichischen Philosophie diskutiert worden sind: „Ist die Referenz des Denkens durch Bezug auf die Referenz der Sprache zu verstehen, oder ist die Referenz der Sprache durch Bezug auf die Referenz des Denkens zu verstehen?“ Es war die Ansicht Brentanos, daß wenn ein *Wort* ein gewisses Ding bezeichnet, so deshalb, weil das Wort fähig ist, *Gedanken* über dieses Ding auszudrücken. Wir können sagen, er verteidigte das Prinzip des Primats des Intentionalen.

Bis vor kurzem haben die meisten der Sprachphilosophen das entgegengesetzte Prinzip akzeptiert – „das Prinzip des Primats des Semantischen“. Dieses Prinzip besagt: Wenn eine Person imstande ist, über ein gewisses Ding zu denken, so deshalb, weil sie *eine Sprache* hat – eine Sprache, in welcher es Wörter gibt, die dieses Ding bezeichnen. Und wie kommt es, daß die Wörter so etwas leisten können? Manchmal lautet die Antwort: „Sie können es einfach – und das ist das Ende“. Ein andermal hört man eine ziemlich skizzenhafte Geschichte über kausale Ketten. Aber keine solche kausale Geschichte ist bei Wittgenstein zu finden.

Was ist sein Standpunkt in Bezug auf die These, daß das Denken auf Sprache beruht?

In den *Philosophischen Untersuchungen* erfaßt er das Problem in einer einzigen einfachen Frage – einer jener merkwürdigen Fragen, die wir nur bei Wittgenstein finden können. Sie lautet, „Was macht meine Vorstellung von ihm zu einer Vorstellung von *ihm*?“ (S. 177) Jene Philosophen, die das Primat des Semantischen verteidigen, würden vielleicht sagen· „Was meine Vorstellung von ihm zu einer Vorstellung von *ihm* macht, ist die Tatsache, daß ich einen inneren Satz behauptet habe – einen Satz, in welchem Wörter, die ihn bezeichnen, verwendet werden.“ Aber Wittgenstein hat gesehen, daß man die gleiche Frage, welche er bezüglich der Vorstellung gestellt hat, auch in Bezug auf den *inneren Satz* stellen kann: „Was macht diesen *Satz* zu einem Satz, der von *ihm* handelt?“ Die Hinwendung zur Sprache dient nur dazu, die ursprüngliche Frage zu verschieben.

Wittgenstein sagt, wenn du wissen willst, an wen ein Mensch denkt, dann brauchst du ihn nur zu fragen. „Und seine Antwort würde entscheiden. Sie würde uns eine *Intention* mitteilen. (*LS* 318, s. *PU* 177). Es ist hier keine Rede von kausalen Ketten oder von dem, was man in den USA „inner systems of representation“ nennt.

Vielleicht habe ich den intentionalen Aspekt Wittgensteins übertrieben. Ich kann mir dessen nicht ganz sicher sein. Aber darüber habe ich volle Gewißheit: Die österreichische Tradition, die mit Bernard Bolzano anfängt und mit Franz Brentano weitergeht und die dann zu Ludwig Wittgenstein führt, ist eine der ertragreichsten in der Geschichte der modernen Philosophie.

* * *

OPENING ADDRESS

Roderick M. Chisholm
Brown University, Providence, R. I.

I thank the program committee for giving me the honor of presenting these opening remarks.

There are many different ways in which we may view the work of a great philosopher. As Wittgenstein might say, we can see it under many different aspects. I will speak this morning of just one of the aspects under which we may view the work of Wittgenstein. This way of looking at what he has done is historically important; I think it may throw light upon his philosophy; and, so far as I know, it has not been discussed in detail. It is this: we view Wittgenstein's work as continuing that of Franz Brentano and his school.

Brentano once wrote an essay on Thomas Reid and divided it into two parts: (1) "was an Reid zu tadeln ist" and (2) "was an Reid zu loben ist." We could not prevent him from writing a long section entitled "was an Wittgenstein zu tadeln ist." But we would have no difficulty in persuading him to write a longer section entitled "was an Wittgenstein zu loben ist."

To see Wittgenstein from the perspective I have in mind, we would do well to begin with his *Remarks on Colour*. His little book, like Brentano's *Untersuchungen zur Sinnespsychologie*, is a treatise about the objects of sensation. Investigation of this area was a feature that was common to the Brentano school and to Austrian positivism as instanced in the work of Mach and Schlick and the early Carnap. It was given up by the positivists when they turned to physicalism, but it was never given up by Wittgenstein. What Wittgenstein and Brentano say about colors would not be acceptable to a philosopher who professed not to know what sensation is. The questions they discuss have nothing to do with physics or even with physical things. As Wittgenstein puts it, the answer to any such question "cannot be a physicalistic statement" (p. 35).

His work on colors, Wittgenstein said, is "not a *theory* of colors (neither a physiological nor a psychological theory), but it is *the logic of the color concept*" (p. 15). Certainly most of Wittgenstein's work is concerned with concepts. Often, when he professes to be talking about *language*, he is really talking about the logic of certain *concepts*. Indeed, where he uses the expression "language game," he could have said, with equal justification, "*conceptual* game."

Wittgenstein also says in the book on colors: "The temptation is very great to believe in a phenomenology, a halfway place between science and logic" (p. 35). Did he succumb to this temptation? He says: "To be sure, there is no phenomenology, but there are phenomenological problems" (22). And these are the problems that he discusses.

The *Philosophische Untersuchungen* is, indeed, an epoch-making contribution to what Brentano had called "descriptive psychology." It extends our knowledge of intentional phenomena to a degree that Brentano's immediate followers had never been able to do. To see this one has only to consider what Wittgenstein tells us there about intentional phenomena—and in particular, about knowing, believing, perceiving and willing. And when he comes to talk about "the dawning of an aspect" he contributes profoundly to our understanding of the concept of *noticing*—a concept that plays a basic role in Brentano's descriptive psychology.

This matter of "the dawning of an aspect" is itself worth singling out. For there is a most striking resemblance between what Wittgenstein does here and what Brentano had done in connection with our experience of temporal relations—what he called *Proteraesthese*.

Other philosophers had noted that our awareness of temporal relations presents us with puzzling philosophical questions. For example: "If you perceive one thing as being prior to another, then you must perceive the first thing as being past. But if a thing is now in the past and no longer in the present, how can you *perceive* it now as *being in the past*?" It was Bren-

tano, I would say, who first saw the full significance of these questions. He saw that they require the re-working of fundamental psychological and philosophical concepts. The result was a theory of the consciousness of time—a contribution which, for good or ill, was to lead to the work of Heidegger and his followers.

Compare, then, what Wittgenstein says about the dawning of an aspect. Psychologists had known that the example of the duck-rabbit presents us with puzzling questions. One was this: "When you cease to see it as a duck and come to see it as a rabbit, what changes—the *object* of sensation or the *act* of sensation?" It was Wittgenstein who first saw the full significance of these questions. And what he has to say about them, it seems to me, has not yet fully been appreciated. Of the causes of the phenomenon, Wittgenstein says: they "are of interest to psychologists, but not to me" (*LW* 434). Here, too, he is discussing descriptive psychology.

There are, of course, important differences between Brentano and Wittgenstein and we should not underestimate them. But we should not exaggerate them either.

I am thinking, for example, of the theory of knowledge—what members of the Brentano school called "the theory of the evident." What Wittgenstein says about evidence—especially in *On Certainty*—may be thought of as supplementing rather than contradicting the doctrine of Brentano.

Wittgenstein does not hesitate to say many of the same things that Brentano does about our immediate experience. Thus he writes: "Certainly one cannot be deceived about immediate experience." But then, unlike Brentano, he adds: "but not because it is certain." (*LW* 187). Where Brentano says "I know" and "It is evident", Wittgenstein says: "It makes no sense to say 'I know' or to say 'It is evident.'" Yet it is difficult to escape the conclusion that the dispute is only verbal. The criteria by means of which Wittgenstein decides that "I know" and "It is certain" are *senseless* are precisely Brentano's criteria for saying that "I know" and "It is certain" are *true*. If a person is in pain, he is not likely to *say* "It is evident to me that I am in pain"; but he is in a position to *use* the fact that he is in pain *as evidence*.

Consider this dialogue:

"Evidently I have not yet recovered."

"How do you know that?"

"I'm still in pain".

Whether or not we say of the person that he *knows* that he is in pain, we can say that he *uses* this fact as evidence for the hypothesis that he has not yet recovered. It would seem that, where the one philosopher concludes, in effect, that the prime mover is unmoved, the other concludes that he moves himself. But in each case the reasons are the same.

It is true that, where Brentano emphasizes *thought*, Wittgenstein emphasizes *language*. Brentano suggests that the principles of logic are to be understood by reference to the rules of thought. And Wittgenstein suggests that they are to be understood by reference to the rules of language. If the one philosopher can be accused of "psychologism", the other can be accused, with equal right, of "linguisticism." But, I would add, the accusation is not justified in either case. How do we decide what the rules of thinking are or what the rules of language are? Whatever the one philosopher can say the other can also say—*mutatis mutandis*.

Are we to understand language in terms of thinking, or conversely? The question turns upon our answer to the most fundamental of all the questions which have been discussed in Austrian philosophy. "Is the reference of thought to be explained in terms of the reference of language? Or is the reference of language to be explained in terms of the reference of thought?" It was Brentano's view that, if a *word* can be said to designate a certain thing, then it does so because it is capable of expressing *thoughts* that refer to that thing. We might say that he defended the principle of the primacy of the intentional.

Until fairly recently, most philosophers of language seemed to accept the contrary principle—what might be called the principle of the primacy of the semantical: if a person is able to think about a certain thing, then he does so in virtue of having a language in which there are words that designate that thing. And how do the words do it? Sometimes the answer is, "They

just do." At other times, one hears a rather sketchy story about causal chains. But no such causal story is to be found in Wittgenstein.

How does he stand with respect to the thesis that thought is based upon language?

In the *Philosophische Untersuchungen*, he puts the problem in one simple question—one of those remarkable questions that one can find only in Wittgenstein. It is this: "What makes my idea of him an idea of *him*?" (p. 177) Those philosophers who affirm the primacy of the semantical would say something like this: "What makes my idea of him an idea of him is the fact that I have asserted an inner sentence in which words are used that refer to him." But Wittgenstein saw that you could raise the same question about the inner *sentence* that he had raised about the idea. "What makes this sentence a sentence about *him*?" The move to language simply transfers the original question.

Wittgenstein says that, if you want to know whom he is thinking of, just ask him. "And his answer would be decisive. It would tell us about an *intention*" (*LW* 318; see *PI* p. 177). There is no talk here about causal chains or "inner systems of representation."

Perhaps I have exaggerated the intentional aspect of Wittgenstein. I cannot be at all certain about this. But I am certain of one thing. The Austrian tradition that begins with Bernard Bolzano and continues with Franz Brentano and then leads on to Wittgenstein is one of the most productive in the history of modern philosophy.

* * *

1. Intentionalität

1. Intentionality

DER INTENTIONALE UND DER REALE GEGENSTAND*

Guido Küng
Universität Freiburg, Schweiz

In der Phänomenologie spielt der Begriff des intentionalen Gegenstandes eine zentrale Rolle. Franz Brentano war von der Auffassung ausgegangen, daß alle psychischen Phänomene sich durch das Merkmal der intentionalen Inexistenz eines Gegenstandes auszeichneten; doch kam er schließlich zur Überzeugung, daß die Annahme solcher in-existierender Gegenstände nicht sinnvoll sei. Husserls Auffassungen entwickelten sich dagegen in gerade entgegengesetzter Richtung: nachdem er anfänglich behauptet hatte, die Phänomenologie solle nur noetische Phänomenologie sein, kam er schließlich zur Überzeugung, daß die Analyse der geistigen Akte nur am Leitfaden der Analyse ihrer intentionalen Korrelate, nämlich der neomatischen Gegenstände, möglich sei.

Der Streit um die intentionalen Gegenstände ist jedoch nicht etwa nur von historischem Interesse; und die Unterscheidung zwischen realem und intentionalem Gegenstand ist mehr als nur ein Übungsfeld für scholastische Spitzfindigkeiten. Es geht hier vielmehr um ein Problem, das für das Wesen des Verhältnisses von Geist und Wirklichkeit zentral ist, und dessen Behandlung heute auch in den Diskussionen der analytischen Philosophie, Logik und Linguistik in den Vordergrund gerückt ist.

In meinem Aufsatz möchte ich zunächst, um die Frage des intentionalen Gegenstandes etwas besser in den Griff zu bekommen, die möglichen Ansichten in einen klassifizierenden Rahmen einordnen. Ich werde deshalb in einem ersten Teil die Vor- und Nachteile von drei allgemeinen Theorien des intentionalen Gegenstandes besprechen, die ich die Transzendenz-Theorie, die Immanenz-Theorie und die Theorie der zwei Objekte nennen will, je nachdem sie den intentionalen Gegenstand als bewußtseinstranszendent oder bewußtseinsimmanent auffaßt, oder für *einen* Akt gleichzeitig zwei intentionale Gegenstände, einen transzendenten *und* einen immanenten, zuläßt. Es wird sich herausstellen, daß keine dieser drei Theorien befriedigt, daß jedoch noch eine vierte Theorie möglich ist, die sich als Meinongsche Theorie bezeichnen läßt. Im dritten Teil meines Aufsatzes werde ich auf diese vierte Theorie näher eingehen und deren Schwierigkeiten und Komplikationen etwas ausführlicher unter die Lupe nehmen.

	Wahrnehmung	Halluzination
Transzendenz-Theorie	Der Akt hat *ein* Objekt: nämlich einen transzendenten Gegenstand, ein Ding in der realen Außenwelt.	Der Akt hat *kein* Objekt.
Immanenz-Theorie	Der Akt hat *ein* Objekt: nämlich einen immanenten Gegenstand, das *Noema* eines Dinges der realen Außenwelt.	Der Akt hat *ein* Objekt: nämlich einen immanenten Gegenstand, das *Noema* eines Dinges, das nicht existiert.
Theorie der zwei Objekte	Der Akt hat *zwei* Objekte: einen immanenten und einen transzendenten Gegenstand.	Der Akt hat *nur ein* Objekt: einen immanenten Gegenstand.

1. *Drei klassische Theorien*

Jeder psychische Akt scheint intentional auf einen Gegenstand gerichtet zu sein. Doch ist es durchaus möglich, daß in Wirklichkeit ein entsprechender Gegenstand gar nicht existiert. Es ist deshalb nicht unmittelbar einsichtig, was unter dem intentionalen Gegenstand eines Aktes zu verstehen ist. Anhand des Gegenstandes von beispielweise Wahrnehmung und Halluzination[1] läßt sich leicht ein Schema von drei verschiedenen Auffassungen zusammenstellen (vergleiche die vorangehende Seite).

1.1 Jede dieser drei allgemeinen Theorien hat ihre Vor- und Nachteile

Die *Transzendenz-Theorie* geht von der Überzeugung aus, daß unser Bewußtsein fähig ist, bewußtseinstranszendente Gegenstände zu erkennen. Sie identifiziert deshalb den intentionalen Gegenstand einer Wahrnehmung mit dem wahrgenommenen Ding in der realen Außenwelt. Im Falle einer Halluzination gibt es jedoch in der realen Außenwelt kein entsprechendes Ding. Die Theorie ist deshalb zur Aussage gezwungen, in einem solchen Falle habe der Akt streng genommen gar kein Objekt. Diese Behauptung ist jedoch unbefriedigend, denn vom phänomenologischen Standpunkt aus scheint es sehr wohl möglich, auch im Falle einer Halluzination ein entsprechendes Objekt, nämlich den halluzinierten Gegenstand, zu beschreiben. Und selbst im Falle einer echten Wahrnehmung hat diese Theorie den Nachteil, daß sie die Tatsache nicht zu berücksichtigen scheint, daß zwei Menschen, die dasselbe Ding betrachten, nie genau dasselbe sehen.

Die *Immanenz-Theorie* vermeidet diese Nachteile der Transzendenz-Theorie. Denn sie identifiziert den intentionalen Gegenstand mit genau dem, was das Subjekt sieht (oder zu sehen meint). Hier ist also der intentionale Gegenstand nicht das Ding in der realen Außenwelt, sondern das Ding-so-wie-es-vom-Subjekt-gesehen-wird. Diese Theorie ist sich bewußt, daß unser Erfassen der Dinge immer subjektiv ist. Da sie aber behauptet, das was wir sehen (oder zu sehen meinen), nämlich der intentionale Gegenstand, dürfe nicht mit dem Ding in der realen Außenwelt verwechselt werden, so hat sie Mühe anzuerkennen, was die Transzendenz-Theorie zum Vornherein als gegeben annimmt, nämlich daß unser Bewußtsein die transzendente Wirklichkeit erfassen kann. Die Immanenz-Theorie führt so zur Behauptung, daß das reale Ding-an-sich unsichtbar sei; und von da gelangt man leicht zur idealistischen Schlußfolgerung, daß die Annahme eines unerkennbaren Dinges-an-sich überflüssig und unsinnig sei, daß es gar keine Dinge-an-sich gebe.

In dieser Lage scheint es angebracht, die *Theorie der zwei Objekte* anzunehmen, da diese die Nachteile sowohl der Transzendenz- als auch der Immanenz-Theorie auszuschließen und die Vorteile dieser beiden Theorien in sich zu vereinigen scheint. Doch hat die Theorie der zwei Objekte ihrerseits ihre Nachteile. Wie Husserl in seiner Besprechung des Buches von Twardowski *Zur Lehre vom Inhalt und Gegenstand der Vorstellungen* angemerkt hat,[2] kann ein und derselbe Akt nicht gleichzeitig *zwei* Objekte haben. (Natürlich können wir z.B. auf einen Blick zwei Schuhe sehen, aber es geht hier natürlich nicht um diese trivialen Fälle der Erkenntnis mehrerer *nebeneinander* stehender Dinge.) Wir sind mit einer Alternative konfrontiert: *entweder* ist das Bewußtsein fähig, ein Ding der realen Außenwelt zu erfassen, und dann müssen wir sagen, daß der Wahrnehmungsakt in diesem Ding der Außenwelt terminiert; *oder* das Bewußtsein ist nicht fähig, wirklich über sich hinauszugehen, und dann müssen wir sagen, daß der Wahrnehmungsakt in einem immanenten Gegenstand terminiert. Ein und derselbe Akt kann nicht gleichzeitig an zwei „Orten" terminieren. Ein Akt, der einen transzendenten Gegenstand erfaßt und ein Akt, der einen immanenten Gegenstand erfaßt beinhalten zwei verschiedene Einstellungen und müssen demnach zwei verschiedene Akte sein.

1.2 Natürlich versucht jede der drei Theorien ihre Nachteile zu überwinden

In der *Transzendenz-Theorie* kann der subjektive Aspekt dessen, was wir sehen (oder zu sehen meinen) dem Akt zugeteilt werden, man kann sagen, er gehöre zu der Wie-Beschaffenheit des Aktes. Sprachlich heißt dies, daß statt Namen von immannenten Gegenständen

adverbiale Ausdrücke zu verwenden sind. Zum Beispiel, statt zu berichten „Ich «sah» rosarote Elefanten" kann man sagen „Ich «sah» rosarot-elefantlich". Diese sogenannte adveriable Theorie ist in der analytischen Philosophie wohl bekannt.[3] Aber auch Husserl hatte eine ähnliche Auffassund vertreten als er in den *Logischen Untersuchungen* einen Bedeutungsbegriff entwickelte, der er später als ‚phansische Bedeutung' kennzeichnete.[4]

Doch Husserl gab sich schon bald Rechenschaft darüber, daß diese Theorie mit der Art wie wir tatsächlich die Bedeutung, d.h. die Richtung unserer Akte, feststellen nicht übereinstimmt. Die adverbiale Theorie schreibt den subjektiven Aspekt dessen, was wir sehen, dem Akt zu, und also sollten wir diesen subjektiven Aspekt nur durch eine Reflektion auf eben diesen Akt erfassen können. Tatsächlich erkennen wir diesen subjektiven Aspekt jedoch „direkt", ohne Reflektion: wir wissen „direkt" wie wir etwas sehen (oder zu sehen meinen). Das Ding-so-wie-ich-es-sehe steht sozusagen „vor" unseren Augen. Die adverbiale Theorie kann dieser „Objektivität" des subjektiven Aspektes insoweit Rechnung tragen, als sie eine passive statt eine aktive Redeweise verwendet. Zum Beispiel, statt „Ich sah rosarot-elefantlich" kann man sagen „Mir erschien-es rosarot-elefantlich" („I was appeared to pink-elephantly"). Doch trägt meiner Meinung nach auch diese Redeweise der Tatsache noch nicht genügend Rechnung, daß der Halluzierende etwas „vor" sich sieht, das in der realen Welt nicht vor ihm steht.

Die *Immanenz-Theorie* hat die entgegengesetzte Aufgabe in ihrem Rahmen einen Ersatz für den transzendenten Gegenstand zu finden. Sie versucht dies dadurch zu erreichen, daß sie den Begriff des immanenten Gegenstandes weiter differenziert. Und zwar beruft sie sich auf die Feststellung, daß es Akte des Identifizierens gibt, in denen verschiedenartige intentionale Gegenstände als ein und derselbe Gegenstand identifiziert werden. So kann ich zum Beispiel den Morgenstern, den ich heute sehe, mit dem Abendstern, den ich gestern gesehen habe, identifizieren. Ich habe in diesem Falle einen Gegenstand vor mir, auf den zwei Kennzeichnungen zutreffen. Die Immanenz-Theorie kann sagen, daß in einem solchen Fall die zwei immanenten Gegenstände, die den zwei Kennzeichnungen entsprechen, für mich denselben „Pol" haben und in diesem Sinne für mich derselbe Gegenstand sind. Bei genauerem Überlegen zeigt es sich jedoch, daß es nicht sehr klar ist, was dieser „identische Pol" sein soll, der „*in*" zwei immanenten Gegenständen vorkommen soll. Vor allem aber ist einzuwenden, daß die Redeweise von dem identischen Pol dem Begriff des transzendenten Gegenstandes nicht gleichwertig ist. Denn Akte des Identifizierens können auch bei Halluzinationen auftreten, wo kein entsprechender transzendenter Gegenstand vorhanden ist. Der wesentliche Unterschied zwischen dem identischen Pol zweier immanenter Gegenstände und dem transzendenten Gegenstand verschwindet nur bei Einnahme eines idealistischen Standpunktes, welcher den transzendenten Gegenstand als eine metaphysische Illusion auffaßt.

Die *Theorie der zwei Objekte* kann nicht, wie die zwei anderen Theorien, dem gegen sie erhobenen Einwand mit *zusätzlichen* Ausführungen begegnen. Sie muß vielmehr ihren Standpunkt *revidieren* und zugeben, daß nur einer der zwei Gegenstände, nämlich der transzendente Gegenstand, das Objekt ist, in dem die Intention terminiert. Wir erhalten somit eine spezielle Version der Transzendenz-Theorie. Der sogenannte immanente Gegenstand darf nun nicht mehr als ein Objekt im strikten Sinne, als ein *quod cognoscitur*, sondern nur noch als ein *medium quo* der Erkenntnis aufgefaßt werden. Husserl scheint eben diesen Begriff eines Pseudo-Objekts,das nur ein *medium quo* ist, entwickelt zu haben als er von der neotischen zu einer explizit neotisch-neomatischen Phänomenologie überging. Zuerst hatte er für dieses *medium quo* die Bezeichnung ‚ontische Bedeutung'[5] verwendet, erst einige Jahre später tritt die Bezeichnung ‚noema' auf. Der Ausdruck ‚ontische Bedeutung' ist wichtig, denn wenn die ontische Bedeutung wirklich eine Bedeutung, d.h. ein Sinn, ist, dann können wir von Frege lernen, daß diese ontische Bedeutung nicht das Referenzobjekt unserer Intention sein kann. Es muß sich in der Tat um ein *medium quo* handeln.

Die Ansicht, daß das Husserlsche Noema als ein Sinn aufzufassen sei, ist besonders von D. Føllesdal und seinen Schülern betont und mit der zeitgenössischen Semantik der möglichen Welten in Beziehung gesetzt worden, vgl. Smith und McIntyre (1982).

Wie ich andernorts ausgeführt habe, vgl. Küng (1975), S. 67, hat es noch weitere Vorteile, wenn die Beziehung zwischen immanentem und transzendentem Gegenstand nach dem Modell von Sinn und Referenzobjekt und nicht nach dem Modell von Wirkung und Ursache oder dem Modell von Zeichen und Bezeichnetem aufgefaßt wird. Der Sinnbegriff kann uns nämlich davor bewahren, das immanente Pseudoobjekt allzu sehr zu reifizieren und in den Repräsentationalismus der Immanenztheorie zu verfallen, wo der immanente Gegenstand als ein Bild oder Zeichen eines unsichtbaren transzendenten Gegenstandes aufgefaßt wird, von dem er verursacht sein soll. Eine Bedeutung, d.h. ein Sinn, ist im Gegensatz zu einem Bild oder Zeichen kein Ding, das ein anderes Ding verdecken kann, sondern sie ist, wie schon gesagt, ein *medium quo*, das seinem Wesen nach auf ein Referenzobjekt hin „durchsichtig" sein kann.

Doch die Auffassung, daß der immanente Gegenstand eine Bedeutung, ein *medium quo* und nicht das *quod* sei, das wir sehen (oder zu sehen meinen), löst nicht alle Probleme. Denn, indem wird die Theorie der zwei Objekte in eine Variante der Transzendenz-Theorie verwandeln, sind wir erneut mit dem Einwand gegen die Transzendenz-Theorie konfrontiert. Wenn nämlich der immanente Gegenstand in ein *medium quo* verwandelt wird, dann kann dieses *medium quo* nicht mehr das Objekt sein, das das Subjekt einer Halluzination zu sehen meint; dann ist dieses *medium quo* nur noch etwas, durch das „hindurch" dieses Subjekt ein *anderes*, nämlich einen Gegenstand, der nicht existiert, sieht. Das Vorhandensein einer immanenten ontischen Bedeutung kann die Schwierigkeit nicht lösen, die bei der Transzendenz-Theorie im Falle des Nichtexistierens des transzendenten Gegenstandes auftritt.

2. *Eine vierte Theorie*

Es scheint zunächst, daß kein Ausweg aus den genannten Schwierigkeiten möglich sei, da die behandelten drei Theorien das gesamte Spektrum der in der Terminologie von immanentem und transzendentem Gegenstand ausdrückbaren Möglichkeiten abzudecken scheinen. Doch wie wir sogleich sehen werden, kann uns ein genaueres Nachdenken über den Begriff des immanenten Gegenstandes weiterhelfen.

Denn was ist unter der Bewußtseinsimmanenz des immanenten Gegenstandes eigentlich zu verstehen? Sowohl Brentano wie auch Husserl betonen beide, daß der immanente Gegenstand nicht als realer (oder reeller) Bestandteil des Bewußtseins aufgefaßt werden kann, daß er nicht im selben Sinne mental ist wie der Bewußtseinsakt. Wie schon erwähnt, hat Brentano beim Nachdenken über diesen Punkt es schließlich aufgegeben, die Redeweise von dem intentionalen In-existieren eines immanenten Gegenstandes wörtlich zu nehmen.

Doch ein erstes läßt sich bezüglich der Immanenz des immanenten Gegenstandes mit Sicherheit festhalten. Die Behauptung, daß ein Gegenstand immanent ist, besagt sicher soviel, daß dieser Gegenstand vorhanden sein kann, auch wenn der transzendente Gegenstand nicht existiert. Die Reduktion des immanenten Gegenstandes auf eine ontische Bedeutung (einen Sinn) trägt dieser These und ebenso der festgestellten Irrealität des immanenten Gegenstandes Rechnung, denn eine Bedeutung ist in der Tat nichts Psychisches und sie kann vorhanden sein, auch wo das entsprechende Referenzobjekt fehlt.

Aber die Bedeutung eines Aktes kann nicht zugleich dessen Objekt sein, und es ist genau dieser Umstand, der uns stört. Denn er ist dafür verantwortlich, daß im Falle der Halluzination die Transzendenz-Theorie nichts enthält, das dem halluzinierten Gegenstand entspricht. Wir müssen deshalb versuchen, trotzdem an der These festzuhalten, daß der immanente Gegenstand Objekt des Bewußtseinsaktes ist. Gleichzeitig wollen wir aber auch die Überzeugung nicht aufgeben, wonach unsere Erkenntnis fähig ist, die transzendente Realität zu erfassen, das heißt wir wollen auch an der These festhalten, daß im Falle einer Wahrnehmung der transzendente Gegenstand das Objekt ist, in dem die Erkenntnis terminisiert. Beide Thesen zugleich lassen sich aber – wenn man die Theorie der zwei Objekte ablehnt – nur unter einer Bedingung festhalten: der *immanente Gegenstand muß in den Fällen, wo der transzendente Gegenstand existiert, mit dem transzendenten Gegenstand identisch sein!* Wir müssen also untersuchen, ob die folgende vierte Theorie des intentionalen Gegenstandes möglich ist:

	Wahrnehmung	Halluzination
Vierte Theorie	Der Akt hat *ein* Objekt: nämlich einen immmanenten Gegenstand, der mit dem transzendenten Gegenstand *identisch* ist.	Der Akt hat *ein* Objekt: nämlich einen immanenten Gegenstand.

Ist diese vierte Theorie nicht ein Widersinn? Wie soll ein immanenter Gegenstand mit einem transzendenten Gegenstand, d.h. mit einem Ding der realen Außenwelt identisch sein können?

Nun, dies hängt von der Bedeutung ab, die wir dem Wort ‚immanent' geben. Wir haben schon gesehen, daß das Adjektiv ‚immanent' nicht besagt, daß der immanente Gegenstand ein realer Bestandteil des Bewußtseins sei. Ebenso ist es falsch, das Bewußtsein als eine Art Schachtel aufzufassen. ‚Im Bewußtsein immanent' kann also auch nicht soviel wie ‚im Bewußtsein eingeschlossen' heißen. Die Immanenz des immanenten Gegenstandes bedeutet vielmehr nichts anderes als, daß *der Gegenstand vom Innern eines Subjektes her vermeint (d.h. intendiert) ist*. Ein Gegenstand der in diesem Sinne immanent ist, kann sehr wohl zugleich transzendent sein, denn ein Ding, das vom Innern eines Subjekts her vermeint ist, kann sehr wohl zugleich ein Ding der realen Außenwelt sein. Aber auch die früher formulierte These, wonach ein immanenter Gegenstand selbst dann vorhanden sein kann, wenn der transzendente Gegenstand fehlt, stimmt: es kommt oft vor, daß ein Ding, das vom Innern eines Subjektes her vermeint ist, in der realen Außenwelt nicht existiert. Übrigens stimmt diese Charakterisierung des immanenten Gegenstandes auch mit dem überein, was der Halluzinierende erlebt: dieser berichtet z.B. nicht, er sehe in seinem Innern, in einem Akt der Reflektion einen rosaroten Elefanten, sondern er berichtet vielmehr, daß er dieses „Tier" irgendwie „vor sich" sehe.

Vielleicht ist der Ausdruck ‚immanenter Gegenstand' so sehr mit unhaltbaren Auffassungen vorbelastet, daß es, um Mißverständnisse zu vermeiden, klüger ist, für die Bedeutung ‚vom Innern eines Subjektes her vermeinter Gegenstand' einen anderen Ausdruck zu verwenden. Dies bietet keine Schwierigkeiten, da uns ja noch der Ausdruck ‚intentionaler Gegenstand' zur Verfügung steht und man ‚intentional' ohne weiteres mit ‚vom Innern eines Subjektes her vermeint' gleichsetzen kann.

3. *Die Schwierigkeiten der vierten Theorie*

Schon ein kurzes Nachdenken zeigt, daß auch die vierte Theorie ihre ganz besonderen Schwierigkeiten hat.

3.1 In den Fällen, wo wie bei der Halluzination zwar ein intentionaler, d.h. vom Innern her vermeinter Gegenstand beschrieben werden kann, aber in der realen Außenwelt dieser Gegenstand nicht vorhanden ist, stellt sich ein besonderes Problem. Soll man sagen, daß in diesen Fällen der von innen her vermeinte Gegenstand existiert oder soll man sagen, daß er nicht existiert? Für das erstere spricht die Tatsache, daß ein Gegenstand vermeint ist, für das letztere die Tatsache, daß dieser vermeinte Gegenstand in Wirklichkeit nicht existiert. Offensichtlich sind wir hier mit dem Meinongschen Problem der nicht-existierenden Gegenstände konfrontiert: haben diese Gegenstände irgendeine Art von Existenz oder haben sie absolut keine Existenz?

Da wir dazu gelangt sind, das Problem genau so zu stellen, wie es sich Meinong gestellt hat, so ist es nicht erstaunlich, daß Meinongs Antwort uns als die natürlichste erscheint: es ist wahr, daß in diesen Fällen etwas vermeint ist, aber es ist ebenso wahr, daß das, was vermeint ist, nicht existiert. Nehmen wir wieder das Beispiel mit den halluzinierten rosaroten Elefanten. Etwas ganz Bestimmtes ist vermeint, nämlich rosarote Elefanten und nicht etwa grüne

Pferde oder Präsident Mitterand. Aber diese rosaroten Elefanten existieren nicht, sie sind bloß vermeint. Tatsächlich ist es für sie durchaus genug, daß sie vermeint sind, denn das Vermeintsein genügt, damit etwas beschrieben werden kann. Als Phänomenologen können wir genau beschreiben, *was* wir vermeinen und *wie* wie es vermeinen, unabhängig davon, ob das Vermeinte nun auch existiert oder nicht. Wir können sogar die *Weise* beschreiben, in der ein vermeinter Gegenstand unserer Meinung nach existiert (ob wir ihm z.B. eine reale oder eine ideale Existenzweise zuschreiben); und wir können beschreiben, *wie überzeugt* wir von seiner Existenz sind (ob wir z.B. wirklich meinen, daß der Gegenstand existiert, oder ob wir, wie in einem Akt des Phantasierens, nur so tun, als ob er existiere). Alles dies können wir beschreiben, unabhängig davon, ob der besagte Gegenstand nun tatsächlich existiert oder nicht.

Aber ist es wirklich richtig, von den nicht-existierenden Gegenständen zu sagen, daß sie gar keine Art von Existenz haben? Ist Vermeint-sein nicht auch schon eine schwache Art des Existierens? Denn heißt Vermeint-sein nicht soviel wie ‚als Objekt eines Subjektes existieren'? Sind wir also nicht gezwungen, die paradoxale Feststellung zu machen, daß die nicht-existierenden Gegenstände doch irgendwie existieren?

Ich glaube nicht, daß dieses Argument zwingend ist. Selbst wenn wir die Redeweise zulassen,[6] daß etwas als Objekt eines Subjektes existiert, so folgt aus der Tatsache, daß etwas als Objekt eines Subjektes existiert, noch nicht, daß dieses Etwas existiert. Aus der Tatsache, daß Wilhelm Tell als Statue existiert, folgt ja auch noch nicht, daß Wilhelm Tell existiert oder existiert hat. Die Aussage ‚*a* existiert als Objekt eines Subjektes' ist nur oberflächlicherweise mit einer Aussage wie z.B. ‚*a* existiert als materielles Ding' verwandt. Aus ‚*a* existiert als materielles Ding' folgt ‚*a* existiert', denn daß *a* als materielles Ding existiert, heißt, daß *a* tatsächlich in der Weise eines materiellen Dinges existiert. Wenn man mir aber nur mitteilt, daß *a* als Objekt eines Subjektes existiert, dann weiß ich noch gar nichts über die Seinsweise von *a*, und ich weiß auch nicht, ob *a* existiert.

Natürlich bestehen zwischen dem Als-Objekt-eines-Subjektes-existieren und dem Als-Statue-existieren auch wesentliche Unterschiede.Eine Statue und eine Person, die durch die Statue dargestellt wird, sind nie identisch, während nach unserer vierten Theorie das intentionale Objekt mit dem transzendenten Gegenstand identisch sein muß, wo ein solcher existiert. Ferner gilt, daß wenn Wilhelm Tell als Statue existiert, dann vielleicht nicht Wilhelm Tell, aber sicher eine Statue existiert. Wenn dagegen ein rosaroter Elefant als Objekt eines Subjektes existiert, dann folgt daraus, weder daß ein rosaroter Elefant existiert, noch auch daß ein Objekt existiert.[7] Das Bestehen dieser Unterschiede soll uns nicht erstaunen, denn Ein-intentionaler-Gegenstand-sein und Eine-Statue-sein ist eben trotz des Bestehens einer gewissen Analogie etwas sehr Verschiedenes.

Aus der Tatsache, daß *a* als Objekt eines Subjektes existiert, folgt nicht, daß *a* existiert; es folgt aber auch nicht, daß *a nicht* existiert. Intentionale Gegenstände sind als solche – um mit Meinong zu sprechen – „jenseits" oder „außerhalb" von Sein und Nichtsein.[8] Natürlich ist dieses Jenseits-von-Sein-und-Nichtsein-sein nicht seinerseits eine Seinsweise, eine Weise des Existierens.

3.2 „‚Ein-intentionaler-Gegenstand-sein' heißt soviel wie ‚vom-Innern-eines-Subjektes-hervermeint-sein". Positiv gesagt ist also das Ein-intentionaler-Gegenstand-sein ein bestimmtes Vermeint-sein. Intentionale Akte sind Akte, die in einer bestimmten Richtung, auf einen bestimmten Punkt hin zeigen. Intentionale Gegenstände, d.h. vermeinte Gegenstände, haben deshalb eine gewisse *Identität*.[9]

Es ist eine Tatsache, daß ein Gegenstand in diesem Sinne identifiziert werden kann, bevor wir wissen, ob er überhaupt vorhanden ist, ob er existiert oder nicht. Jedermann kann sich zum Beispiel ein Pferd vorstellen und dann annehmen, daß *dieses* Pferd über ein Hindernis springt und daß *es* stürzt, sich ein Bein bricht, usw. usf. Bei der Beschreibung solcher Fälle ist es völlig normal, hinweisende Artikel und Pronomina zu gebrauchen, welche angeben, daß ständig derselbe Gegenstand vermeint wird. Es ist überdies in der Umgangssprache durchaus möglich, einen in der Phantasie vermeinten Gegenstand mit einem in der realen Wirklichkeit

existierenden Gegenstand zu identifizieren. Es ist z.B. denkbar, daß, während ich nicht glaube, daß das Loch Ness Ungeheuer existiert und ich mir bloß in der Phantasie vorstelle, daß *es* im Loch Ness herumschwimmt, *dieses* Ungeheuer tatsächlich existiert und jemand *es* gesehen hat . . . Dieser Gebrauch der Pronomina und der hinweisende Artikel in der Umgangssprache zeigt, daß die vierte Theorie keineswegs ein künstliches Gebilde ist, sondern sogar sehr genau mit den Redeweisen der natürlichen Sprache harmoniert.

Die intentionalen Gegenstände können in etwa mit geographischen *Koordinationspunkten* verglichen werden,[10] und die Frage, ob sich die intentionalen Gegenstände innerhalb oder außerhalb des Bewußtseins befinden, ist wie die Frage, ob sich die geographischen Koordinationspunkte im Bewußtsein der Geographen oder in der Außenwelt befinden. Man könnte argumentieren, daß es ohne Geographen keine geographischen Koordinationspunkte geben würde, und daß sich diese Punkte folglich im Bewußtsein des Geographen befinden müßten. Aber dieses Argument ist nicht überzeugend. Denn, wenn man einem Geographen sagt, daß an einem bestimmten Koordinatenpunkt ein Schatz vergraben liegt, dann sucht er nicht in seinem Bewußtsein nach diesem Schatz, sondern verifiziert, ob dieser Punkt mit einem bestimmten Ort der Erde identisch sei.

3.3 Doch die Annahme nicht existierender Gegenstände wirft ein weiteres Problem auf, das meiner Ansicht nach den eigentlichen Kern der Problematik der nicht-existierenden Gegenstände bildet: Können nicht-existierende Gegenstände Eigenschaften haben, und falls sie Eigenschaften haben, welcher Art sind diese Eigenschaften?

3.3.1. Nehmen wir den Fall, wo ich mir in reiner Phantasie ein grünes Pferd vorstelle. In diesem Fall ist etwas als ein Pferd und als grün-seiend *vermeint.*[11] Aber kann man auch sagen, daß hier etwas ein Pferd *ist* und grün *ist*? Das fragliche „grüne Pferd" existiert ja nicht, wie kann es da *grün* sein, so wie z.B. ein existierender Baum grün ist?[12] Aber es ist sicher *als-grün-seiend-vermeint.* Von daher scheint es plausibel, den nicht-existierenden Gegenständen statt der gewöhnlichen Eigenschaften nur sogenannte konvers-intentionale Eigenschaften[13] wie Als-ein-Pferd-vermeint-zu-sein und Als-grün-seiend-vermeint-zu-sein zuzuschreiben. Bezüglich der konvers-intentionalen Eigenschaften scheint es noch am ehesten möglich, die erstaunliche These zu akzeptieren, daß sie exemplifiziert sein können, selbst wenn der exemplifizierende Gegenstand ein nicht-existierender Gegenstand ist.[14]

Doch nicht-existierende Gegenstände müssen auch andere als konvers-intentionale Eigenschaften haben. Denn nicht-existierende Gegenstände sind tatsächlich nicht-existierend, sie haben also auch die Eigenschaft nicht-existierend zu sein, und diese Eigenschaft ist keine konvers-intentionale Eigenschaft. Ein Gegenstand kann die Eigenschaft des Nicht-existierend-seins ja sogar dann besitzen, wenn er von einem Subjekt als existierend vermeint ist. Aber das Nicht-existierend-sein ist keine gewöhnliche Eigenschaft, und so können wir weiterhin behaupten, daß die nicht-existierenden Gegenstände sich dadurch auszeichnen, daß sie keine gewöhnlichen Eigenschaften haben.

3.3.2 Doch die These, daß ein vorgestellter Gegenstand, der nicht existiert, nicht die Eigenschaft Grün, sondern nur die Eigenschaft Als-grün-seiend-vermeint-zu-sein haben könne, kann auch in Zweifel gezogen werden. Denn wenn ich mir in meiner Phantasie ein grünes Pferd vorstelle, stelle ich mir dann wirklich etwas vor, das nur als-grün-seiend-vermeint ist? Sagen wir in einem solchen Fall nicht ganz spontan, daß ich mir etwas vorstelle, das *grün ist*?

Meinong vertritt deshalb in Übereinstimmung mit der Umgangssprache, die These, daß im besprochenen Beispiel der vorgestellte Gegenstand ein Pferd ist und grün ist.[15] Die Merkwürdigkeit, daß demnach etwas Nicht-existierendes ein Pferd und grün sein kann, sucht Meinong mit seiner Lehre vom *Sosein* und *Dasein* zu erklären. Das vorgestellte grüne Pferd hat kein Dasein, aber es hat das Sosein eines Pferdes und es hat das Sosein eines grünen Dinges.

Ich kann mir aber auch vorstellen, daß ein grünes Pferd in der realen Außenwelt existiert. Was ich mir vorstelle ist dann nicht nur ein Pferd und grün, es ist auch ein in der realen

Außenwelt Existierendes. Gleichzeitig ist es aber auch ein nichtexistierender Gegenstand. Wie ist das möglich? Offensichtlich muß hier, wenn man nicht zwei Gegenstände annehmen will, die Kopula zwei verschiedene Bedeutungen haben:[16]

(1) Das Vorgestellte ist_1 ein in der realen Außenwelt existierendes grünes Pferd.
(2) Das Vorgestellte ist_2 ein nicht-existierender Gegenstand.

Wie ist aber der Unterschied dieser zwei Kopulas zu verstehen?

Die These, die den nicht-existierenden Gegenständen die gewöhnlichen Eigenschaften abspricht, kann den Unterschied zwischen diesen Kopulas mit Hilfe des weniger mysteriösen Unterschiedes zwischen der Prädikation von konvers-intentionalen Eigenschaften und der Prädikation von nicht konvers-intentionalen Eigenschaften wegerklären:

(3) Das Vorgestellte *ist vermeint als* ein in der realen Außenwelt existierendes grünes Pferd.
(4) Das Vorgestellte *ist* ein nicht-existierender Gegenstand.

In (3) ist die Kopula ‚ist_1' von (1) durch den Ausdruck ‚ist vermeint als' übersetzt worden. Der Ausdruck ‚ist vermeint als' ist zusammengesetzt aus der Kopula ‚ist' und dem Ausdruck ‚vermeint als', der dem Prädikat zuzuzrechnen ist. Es kommt also in (3) und (4) nur noch *eine* Kopula vor. Allerdings handelt es sich, um eine etwas ungewöhnliche Kopula, da sie ja, wie schon die Kopulas ‚ist_1' und ‚ist_2' auch dann eine wahre Aussage bilden kann, wenn der dem Subjekt-Terminus entsprechende Gegenstand nicht existiert. Es genügt für diese Kopula, daß der Subjektgegenstand eine gewisse Identität besitzt.

Daß die These, die den nicht-existierenden Gegenständen die gewöhnlichen Eigenschaften abspricht, die Schwierigkeit mit den zwei Kopulas wegerklären kann, ist der große Vorteil dieser These. Die Alternative scheint also zu lauten: entweder (a) wir schreiben den nicht-existierenden Gegenständen gewöhnliche Eigenschaften zu: dann sind wir in Übereinstimmung mit der Umgangssprache, müssen aber die Schwierigkeit mit den zwei Kopulas in Kauf nehmen; oder (b) wir schreiben den nicht-existierenden Gegenständen statt der gewöhnlichen Eigenschaften nur entsprechende konvers-intentionale Eigenschaften zu: dann sind wir die Schwierigkeit mit den zwei Kopulas los, doch verwerfen wir in diesem Fall gewisse gängige umgangssprachliche Redeweisen.

3.3.3 Um die Alternative intuitiv etwas besser in den Griff zu bekommen, wollen wir einen Fall näher betrachten, der zwischen Wahrnehmung und Fiktion eine Mittelstellung einnimmt, nämlich den Fall einer sich täuschenden Wahrnehmung (delusory perception, misperception)[17], wo ich ein Maultier für ein Pferd halte.

Wenn wir die Analyse des intentionalen Gegenstandes mit Hilfe von konvers-intentionalen Eigenschaften durchführen, dann geht alles ohne Schwierigkeiten. Wir finden dann, daß der intentionale Gegenstand hier etwas ist, das neben anderen die zwei Eigenschaften Als-ein-Pferd-vermeint-zu-sein und Nicht-als-ein-Maultier-vermeint-zu-sein hat. Dieser intentionale Gegenstand ist nach Auffassung der vierten Theorie mit dem transzendenten Gegenstand identisch, der neben anderen die zwei Eigenschaften Ein-Maultier-zu-sein und Nicht-ein-Pferd-zu-sein hat. Es gibt also nach dieser Analyse einen einzigen Gegenstand, der neben anderen die vier folgenden Eigenschaften hat: Ein-Maultier-zu-sein, Nicht-ein-Pferd-zu-sein, Als-ein-Pferd-vermeint-zu-sein, Nicht-als-ein-Maultier-vermeint-zu-sein. Diese vier Eigenschaften sind problemlos kompatibel und der Gegenstand ist ein ganz normales existierendes Maultier.

In Übereinstimmung mit der Umgangssprache kann man aber bei dieser sich täuschenden Wahrnehmung noch von etwas anderem als von diesem Maultier sprechen. Das Maultier ist das, was *effektiv* vor dem Betrachter steht. Man kann aber auch von dem sprechen, was vor dem Betrachter stehen *müßte*, wenn die falsche Behauptung ‚Vor mir steht ein Pferd' wahr

wäre: es müßte dann etwas vor mir stehen, das die Eigenschaft Ein-Pferd-zu-sein hat; auch müßte dieses Etwas natürlich die Eigenschaft des Existierend-seins haben. Da ich während der Täuschung tatsächlich meine, daß die Behauptung ‚Vor mir steht ein Pferd' wahr sei, so kann man sagen, daß während der Täuschung meine Intention nicht nur auf das Maultier, sondern auch auf dieses Pferd gerichtet sei.

Das Pferd, das da sein müßte, wenn die Behauptung ‚Vor mir steht ein Pferd' wahr wäre, kann mit dem Maultier, das effektiv vor mir steht *nicht* identisch sein, da das Maultier kein Pferd ist. Auch ist das Pferd, das da sein müßte, in Wirklichkeit nicht da, es ist ein nicht-existierender Gegenstand. Wie ersichtlich, muß es sich um einen Gegenstand handeln, der die oben erwähnte Schwierigkeit mit den zwei Kopulas mit sich bringt: der besagte Gegenstand ist_1 ein existierendes Pferd und er ist_2 ein nicht-existierender Gegenstand. Natürlich hat dieser nicht-existierende Gegenstand im Augenblick der Täuschung nicht die Eigenschaft, als ein nicht-existierender Gegenstand vermeint zu sein. Im Gegenteil, er ist_2 zu jenem Zeitpunkt als-existierend-vermeint, nämlich als-ein-existierendes-Pferd-vermeint.

Wir scheinen also bei diesem Beispiel einer sich täuschenden Wahrnehmung gleich wie vorher beim Beispiel einer reinen Phantasievorstellung mit einer Alternative von zwei Objekten konfrontiert zu sein. Doch eine etwas sorgfältigere Überlegung zeigt, daß die Alternative eine etwas komplexere Gestalt hat. Sie lautet nämlich nicht: *entweder* nehmen wir den Gegenstand der ersten Analyse, nämlich das Maultier, das die konvers-intentionale Eigenschaft des Als-Pferd-vermeint-seins hat an, *oder* wir nehmen den Gegenstand der zweiten Analyse, das Pferd, das zugleich ein nicht-existierender Gegenstand ist, an. Die Alternative lautet vielmehr: *entweder* nehmen wir nur den Gegenstand der ersten Analyse an, *oder* wir nehmen sowohl den Gegenstand der ersten Analyse als auch den Gegenstand der zweiten Analyse an.

Vom Standpunkt des Sparsamkeitsprinzips der analytischen Philosophie ist es empfehlenswert, bei der Beschreibung der sich täuschenden Wahrnehmung nur auf das keine Schwierigkeiten machende Maultier zu sprechen zu kommen. Vom Standpunkt der Phänomenologie hingegen, scheint der Gegensatz zwischen dem, was vor mir stehen müßte, und dem, was effektiv vor mir steht, für das Wesen der Täuschung besonders charakteristisch zu sein. Der Phänomenologe ist nicht in erster Linie an Sparsamkeit interessiert, sondern er sucht im Gegenteil eine möglichst reichhaltige Ontologie zu entwickeln. Er wird also versuchen beide Gegenstände, sowohl das Maultier als auch das nicht-existierende Pferd in seine Beschreibung einzubauen.

Der Phänomenologe erhält so eine neuartige „Theorie mit zwei Objekten". Es handelt sich dabei nicht etwa um eine Version der früheren Theorie der zwei Objekte, die zwischen dem transzendenten Gegenstand und einem als ‚immanent' bezeichneten Gegenstand-genau-so-wie-er-vermeint-ist unterschied. Denn bei dieser neuartigen Theorie mit zwei Objekten ist ja jedes der zwei Objekte, sowohl das Maultier als auch das nicht-existierende Pferd, ein Objekt im Sinne der vierten Theorie: jedes von ihnen hat sowohl ein gewisses An-sich-sein, das von den Erkenntnisakten des Subjekts unabhängig ist, als auch ein gewisses Vermeint-sein, d.h. gewisse konvers-intentionale Eigenschaften.[18]

3.3.4 Ein Unterschied zwischen dem, was vor mir stehen müßte, und dem, was effektiv vor mir steht, ist dem Phänomenologen noch von einem anderen Beispiel her geläufig. Ein analoger Unterschied tritt nämlich beim Versuch einen kontradiktorischen Gegenstand zu denken auf.

Nehmen wir den Versuch ein rundes Viereck zu denken. In diesem Falle *müßte*, wenn mein Denkversuch gelingen würde, ein Ding vor meinem geistigen Auge stehen, das zugleich rund und viereckig *ist*. Doch die Eigenschaften Rund-zu-sein und Viereckig-zu-sein sind nicht kompatibel, mein Denkversuch kann deshalb unmöglich zu dem erstrebten Ende kommen. Der Gegenstand, der rund und viereckig ist_1, ist_2 ein unmöglicher Gegenstand, der weder existieren noch auch „effektiv gedacht" werden kann. Effektiv steht vor meinem geistigen Auge nur ein Etwas, das die Eigenschaft hat als rund und viereckig *vermeint* zu sein, das aber nicht rund und viereckig *ist*.

Natürlich ist die Unterscheidung zwischen den zwei Objekten beim Beispiel des Denkens eines runden Vierecks schwerer nachzuvollziehen als beim Beispiel der sich täuschenden Wahrnehmung. Bei der sich täuschenden Wahrnehmung war nur eines der beiden Objekte ein nicht-existierender Gegenstand, während das andere Objekt ein handfestes Maultier war. Beim Denken eines runden Vierecks sind beide Objekte nicht-existierende Gegenstände, und überdies ist hier gerade das Etwas, das dem Maultier entspricht und das wie das Maultier logisch unproblematischer ist, ein Etwas, von dem man im Alltag nie spricht. Wir haben deshalb große Mühe, dieses Etwas zu bemerken, obwohl es im Gegensatz zum runden Viereck effektiv vor unserem geistigen Auge steht und obwohl es im Gegensatz zum runden Viereck kein kontradiktorischer Gegenstand ist und mit nur einer Kopula auskommt. Ich meine aber, daß die Analogie zum Maultier der sich täuschenden Wahrnehmung uns helfen kann, dieses ungewohnte Etwas zu erfassen.

3.3.5 Kehren wir nunmehr, nach dem Exkurs über die sich täuschende Wahrnehmung und das Denken unmöglicher Gegenstände, wieder zur reinen Phantasievorstellung eines grünen Pferdes zurück. Wir bemerken sofort, daß es auch hier bei den zwei Objekten um eine Unterscheidung zwischen dem, was vor mir stehen müßte, und dem, was effektiv vor mir steht, geht. Phantasieren heißt ja in einem gewissen Sinne so tun, als ob man etwas Bestimmtes wahrnehmbar vor sich hätte; es ist in einem gewissen Sinne ein Wünschen (aber natürlich nicht ein ernst gemeintes Wünschen), daß etwas wahrnehmbar vor mir stehe. Im gegebenen Fall ist ein grünes Pferd das, was vor mir stehen müßte, wenn der Phantasiewunsch erfüllt sein sollte. Was hingegen effektiv vor meinem geistigen Auge steht, ist ein Etwas, das die Eigenschaft hat als-ein-wahrnehmbar-vor-mir-stehendes-grünes-Pferd-*vermeint-zu-sein*, das aber nicht die Eigenschaft hat, ein-wahrnehmbar-vor-mir-stehendes-grünes-Pferd-*zu-sein*.

3.3.6 Ziehen wir aus dem zu den drei Beispielen Gesagten die Bilanz. In allen diesen Fällen kann man von zwei Objekten sprechen, nämlich von dem, was in einem gewissen Sinne (und in jedem Beispiel wieder in einem etwas anderen Sinne) vor mir stehen *müßte*, und von dem, was in einem gewissen Sinne *effektiv* vor mir steht. Der Umstand, daß die zwei Objekte nicht identisch sind, daß das, was vor mir stehen müßte, und das, was effektiv vor mir steht, auseinanderfällt, besagt, daß in allen diesen Beispielen eine gewisse (aber je wieder anders geartete) Unerfülltheit besteht. Gewöhnlich sind die Phänomenologen mehr am Thema der Erfüllung als am Thema der Unerfülltheit interessiert. Doch ist natürlich das eine dieser Themen immer die Kehrseite des anderen. Die hier entwickelte Unterscheidung von zwei Objekten, die beide Objekte im Sinne der vierten Theorie sind, kann also auch als ein kleiner Beitrag zur Analyse der Erfülltheit aufgefaßt werden. Man kann ja sagen, daß die Erfülltheit neben anderen Merkmalen das Merkmal einer ganz bestimmten Einfachheit, nämlich eben der Abwesenheit jener Unterscheidung, beinhaltet. Tatsächlich ist bei einer täuschungsfreien Wahrnehmung, beim Sich-denken eines möglichen Gegenstandes, beim Wahrnehmen und nicht nur Phantasieren, das, was vor mir stehen soll, und das, was effektiv vor mir steht, nicht mehr unterschieden. Es gibt jeweils nur noch *ein* Objekt, und zwar das logisch problemlose Objekt, und zudem stimmen alle konvers-intentionalen Eigenschaften dieses Objektes mit seinen „gewöhnlichen" Eigenschaften überein. Wenn zum Beispiel ein Maultier täuschungsfrei als ein Maultier wahrgenommen wird, dann hat dieses Maultier sowohl die Eigenschaft Als-ein-Maultier-vermeint-zu-sein als auch die Eigenschaft Ein-Maultier-zu-sein.

In dem vorliegenden Aufsatz habe ich die Unterscheidung der zwei Objekte im Sinne der vierten Theorie nur intuitiv und informell skizziert. Wie ich jedoch nachträglich festgestellt habe, sind diese zwei Objektbegriffe nicht nur formalisier*bar*, sie scheinen zum Teil schon weitgehend formalisiert worden zu sein, allerdings nicht zusammen in *einem* System, sondern getrennt in zwei verschiedenen Systemen. Diese Arbeiten gehen auf zwei verschiedene Anregungen von Ernst Mally zurück, wie die Schwierigkeiten der Meinongschen Gegenstandstheorie gelöst werden könnten. Diese Anregungen sind durch Findlay (1933) einem angelsächsischen Leserkreis vermittelt worden.

Die eine Anregung[19] ist von Meinong selbst als Lösung akzeptiert worden. Sie entspricht ungefähr meiner Unterscheidung der zwei Kopulas. In Meinongs Terminologie ausgedrückt handelt es sich um die Unterscheidung zwischen „konstitutorischen Eigenschaften" (entsprechend der Kopula ‚ist$_1$') und „außerkonstitutorischen Eigenschaften" (entsprechend der Kopula ‚ist$_2$'). Findlay spricht von „nuclear properties" und „extra-nuclear properties". Die ausführlichste Ausarbeitung eines Systems, das auf dieser Anregung basiert, findet sich in Parsons (1980).

Die zweite Anregung Mallys[10] entspricht in gewisser Hinsicht meiner Unterscheidung zwischen dem Haben von konvers-intentionalen Eigenschaften und dem Haben von nicht-konvers-intentionalen Eigenschaften. Mally empfiehlt hier nämlich zwischen dem Bestimmt-sein durch eine Bestimmung und dem Erfüllen einer Bestimmung zu unterscheiden. Er nimmt zum Beispiel an, das sogenannte runde Viereck sei nur durch die Bestimmung des Rund-und-viereckig-seins *bestimmt*, aber es *erfülle* diese Bestimmung nicht (Findlay würde sagen „it is determined by this determination, but it does not satisfy it".) Dies entspricht in etwa meiner Rede von einem Etwas, das nur als rund-und-viereckig-seiend-vermeint ist, aber nicht rund-und-viereckig ist. Doch ist die Entsprechung zwischen dem Mallyschen Bestimmt-sein und meinem Vermeint-sein nur in beschränktem Umfange vorhanden. Zum einen ist das Mallysche Bestimmt-sein unpersönlich verstanden, während ein Vermeint-sein nur vorkommen kann, wenn es eine Person gibt, die vermeint. Zum anderen scheint Mally nur von abstrakten Gegenständen zu sagen, daß sie bestimmt werden: die Bestimmung des Pferd-seins bestimmt nur den abstrakten Gegenstand „ein Pferd" und nicht ein konkretes existierendes Pferd. Trotz dieser Unterschiede besteht aber doch, wenigstens was die abstrakten (d.h. nicht-existierenden) Gegenstände betrifft, eine große Ähnlichkeit zu dem Objektbegriff, den ich als den logisch weniger problematischen bezeichnet habe.

Die zweite Anregung von Mally ist 1983 von Zalta, einem Schüler Parsons', systematisch ausgearbeitet worden. Dabei wurde die Terminologie etwas geändert: statt vom Erfüllen einer Bestimmung spricht Zalta vom Exemplifizieren einer Eigenschaft; und an Stelle der passiven Redeweise, die sagt, daß ein abstrakter Gegenstand *durch eine Bestimmung bestimmt werde*, führt Zalta eine positive Redeweise ein, nach der dieser Gegenstand *eine Eigenschaft enkodiere* (encodes).[21] Zalta nimmt zum Beispiel an, daß es nichts gibt, das die komplexe Eigenschaft Rund-und-viereckig-zu-sein exemplifiziert, daß aber ein gewisser abstrakter Gegenstand diese Eigenschaft enkodiert. Im Vergleich zu meinen eigenen Ausführungen besteht ein gewisser Unterschied, indem Zalta von zwei verschiedenen Beziehungen zwischen Gegenständen und Eigenschaften, dem Exemplifizieren und dem Enkodieren, aber nicht von zwei Arten von Eigenschaften, wie zum Beispiel dem Rot-exemplifizierend-sein (d.h. Rot-sein) und dem Rot-enkodierend-sein, spricht. Doch die letztere Redeweise ließe sich wohl ohne Schwierigkeiten in sein System einbauen.

ANMERKUNGEN

* Eine frühere Fassung dieses Aufsatzes ist in englischer Sprache in der Zeitschrift *Dialectica* Bd. 38 (1984), S. 143–156 erschienen.

[1] Ich verstehe unter Halluzination das, was Smith (1983), S. 97–98 als naive Halluzination bezeichnet. Smith unterscheidet zwischen illusionäre Wahrnehmung (illusory perception: es existiert etwas in der Wirklichkeit, aber es ist nicht so wie es dem wahrnehmenden Subjekt erscheint), sich täuschender Wahrnehmung (delusory perception: es existiert etwas in der Wirklichkeit, es ist nicht so wie es erscheint, aber das wahrnehmende Subjekt meint, es sei so wie es erscheint), naiver Halluzination (naive hallucination: etwas erscheint einer halluzinierenden Person und diese Person meint das Erscheinende existiere in der Wirklichkeit, aber tatsächlich entspricht dem Erscheinenden nichts in

der Wirklichkeit), aufgeklärte Halluzination (hip hallucination: etwas erscheint einer halluzinierenden Person als ob es in der Wirklichkeit existiere, doch die halluzinierende Person nimmt richtigerweise an, daß dem Erscheinenden in der Wirklichkeit nichts entspreche) und neutraler Halluzination (neutral hallucination: etwas erscheint einer halluzinierenden Person, tatsächlich entspricht dem Erscheinenden nichts in der Wirklichkeit, doch die halluzinierende Person läßt die Frage unentschieden, ob dem Erscheinenden in der Wirklichkeit etwas entspreche oder nicht). – Es wird oft behauptet, daß, was die subjektive Erfahrung betrifft, zwischen einer naiven Halluzination und einer echten Wahrnehmung kein Unterschied besteht. Doch scheint mir diese Auffassung nicht richtig zu sein. Ich würde eher sagen, daß die einer naiven Halluzination unterliegende Person die Merkmale, die ihre subjektives Erleben von dem Erleben einer echten Wahrnehmung unterscheiden, nicht bemerkt, daß solche Merkmale jedoch vorhanden sind.

[2] Husserl (1896), S. 351 Fußnote: „Die Vorstellungstätigkeit bewegt sich nicht in doppelter Richtung."

[3] Vgl. z.B. Butchvarov (1980), Tye (1984).

[4] Vgl. Küng (1973). Wie ich aufgewiesen habe, hat Husserl in seinen „Vorlesungen über Bedeutungslehre" vom Sommersemester 1908 (unveröffentlichtes Manuskript F I 5) zwischen phansischer und ontischer Bedeutung unterschieden.

[5] Vgl. Anmerkung 4. Husserl spricht von der ontischen Bedeutung als dem „Gegenstand genommen in der Weise wie er bedeutet oder gedacht ist", und setzt den so verstandenen Gegenstand dem „Gegenstand schlechthin" gegenüber, der im Falle eines falschen Urteils fehlen kann. Wenn man den Ausdruck ‚Bedeutung' mit ‚Sinn' gleichsetzt, dann ist man geneigt die Beziehung zwischen dem „Gegenstand genommen in der Weise wie er bedeutet oder gedacht ist" und dem „Gegenstand schlechthin" als die Beziehung zwischen Sinn und Referenzobjekt aufzufassen. Man kann aber auch versuchen, diese Husserlschen Redeweisen im Sinne der vierten Theorie zu verstehen, auf die ich im Folgenden zu sprechen kommen werde.

[6] Vgl. jedoch Anmerkung 7.

[7] Der Umstand, daß im Rahmen der vierten Theorie aus ‚*a* existiert als Objekt eines Subjektes' nicht gefolgert werden darf ‚Ein Objekt existiert', legt nahe, daß es vielleicht doch sinnvoller ist, hier die Redeweise, daß etwas als Objekt eines Subjektes existiert, zu verbieten.

[8] Vgl. Meinong (1904), S. 494: „man könnte vielleicht sagen: der reine Gegenstand stehe «jenseits von Sein und Nichtsein». Minder ansprechend oder auch minder anspruchsvoll, dafür aber meines Erachtens sonst geeigneter, ließe sich dasselbe etwa so aussprechen: der Gegenstand ist von Natur außerseiend . . ."

[9] Ich bin mir bewußt, daß die Identität intentionaler Gegenstände sehr viele schwierige Probleme aufwirft und spreche denn auch nur von einer „gewissen" Identität. Vgl. z.B. die Ausführungen in Küng (1980), S. 112–113, welche betonen, daß der intendierte Gegenstand oft nicht gegeben, sondern nur „aufgegeben" ist.

[10] Vgl. hierzu Findlay (1933), S.112.

[11] Genau genommen müßte man hier nicht sagen, daß etwas als grün-seiend vermeint wird, sondern nur, daß etwas *vermeint* wird, *als ob es als etwas Grünes sichtbar wäre*. Denn im Gegensatz zu einer Wahrnehmung oder einer naiven Halluzination meinen wir ja im Falle einer reinen Phantasievorstellung nicht wirklich, daß etwas ein grünes Pferd sei; wir tun nur als ob.

[12] Vgl. hierzu Husserl (1894), S. 308 Anmerkung: „. . . was nicht existiert, ‹kann› auch nicht Eigenschaften haben ‹. . .›; sie werden nur als einem Gegenstand zukommend vorgestellt . . ."

[13] Der Ausdruck ‚konvers-intentionale Eigenschaften' wurde von Chisholm (1984) geprägt. – Es ist in diesem Zusammenhang wichtig einzusehen, daß eine Aussage wie ‚Der von mir vermeinte Gegenstand ist grün' zweideutig ist. Diese Aussage kann bedeuten „Der von mir vermeinte Gegenstand ist als grün-seiend vermeint", sie kann aber auch bedeuten „Der von mir vermeinte Gegenstand ist tatsächlich grün" (obwohl ich ihn vielleicht als rot-seiend vermeint habe).

[14] Zalta (1983) spricht davon, daß abstrakte Gegenstände Eigenschaften wie Grün-sein nur „enkodieren" aber nicht exemplifizieren können (vgl. unten Anmerkung 21). Er wirft aber nicht die Frage auf, ob diese abstrakten Gegenstände Eigenschaften wie Grün-enkodierend-sein exemplifizieren.

[15] Vgl. Brentano (1930), S. 88: „. . die Vorstellung eines Pferdes ‹hat› nicht «vorgestelltes Pferd», sondern «Pferd» . . . zum Objekt". – Es ist jedoch zu betonen, daß der Unterschied zwischen der Phantasievorstellung eines Pferdes und der Vorstellung der Phantasievorstellung eines Pferdes auch nur schon in den konvers-intentionalen Eigenschaften zum Ausdruck kommt: im ersten Fall hat etwas die Eigenschaft Als-ob-es-ein-Pferd-wäre-vermeint-zu-sein, im zweiten Fall hat etwas die Eigenschaft Als-ein-in-der-Phantasie-vermeintes$_1$-Pferd-vermeint$_2$-zu-sein

[16] Das Problem einer solchen Doppelstruktur tritt auch bei Philosophen auf, die keine nicht-existierenden Gegenstände annehmen. So unterscheidet z.B. Frege (1884), S. 64 zwischen den Merkmalen und den Eigenschaften von Begriffen; und Ingarden zwischen der Struktur und dem Gehalt von rein intentionalen Gegenständen und von Ideen: vgl. Ingarden (1931), § 20 und Ingarden (1965), § 47 und § 50.

[17] Vgl. Anmerkung 1.

[18] Wir haben in Bezug auf die frühere Theorie der zwei Objekte den Einwand erhoben, daß ein und derselbe Akt nicht gleichzeitig zwei verschiedene Objekte haben könne. Muß dieser Einwand nicht auch bei dieser neuartigen Theorie berücksichtigt werden? Die Lage ist jedoch hier eine besondere,

da das Subjekt im Augenblick der Täuschung von der Verschiedenheit der zwei Objekte nichts weiß. Erst in der Retrospektive kann eine Täuschung als Täuschung ins Blickfeld des Phänomenologen geraten und erst dann wird die Verschiedenheit der zwei Objekte festgestellt. – Im Folgenden (3.3.4) werden wir den Fall des Denkens eines runden Vierecks besprechen. Hier ist unsere gegenwärtige Frage schwieriger zu beantworten. Vielleicht muß man da unterscheiden zwischen dem Akt, der das runde Viereck zu denken versucht, und dem Akt, der feststellt, daß nichts Rundes-und-Viereckiges vor meinem geistigen Auge „erschienen" ist.

[19] Mally (1908), S. 882, Meinong (1915), S. 175–177. Findlay (1933), S. 176, Zalta (1983), S. 10.
[20] Mally (1912), S. 64 und 76, Findlay (1933), S. 110–112 und 182–184, Zalta (1983), S. 11.
[21] Zalta (1983), S. 12.

LITERATUR

Brentano, F., *Wahrheit und Evidenz* (Leipzig 1930).
Butchvarov, P., „Adverbial theory of consciousness", *Midwest Studies in Philosophy* Bd. 5 (1980).
Chisholm, R.M., „What is the problem of objective reference?" *Dialectica* Bd. 38 (1984).
Findlay, J.N., *Meinong's Theory of Objects* (London 1933). Um zwei Kapitel vermehrte Ausgabe unter dem Titel *Meinong's Theory of Objects and Values* (Oxford 1963).
Frege, G., *Grundlagen der Arithmetik* (Breslau 1884).
Husserl, E., „Intentionale Gegenstände" (1894), in: *Husserliana* Bd. 22 (Den Haag 1979).
Husserl, E., Besprechung von K.Twardowski, *Zur Lehre vom Inhalt und Gegenstand der Vorstellungen* (1896), in: *Husserliana* Bd. 22 (Den Haag 1979).
Ingarden, R., *Das literarische Kunstwerk* (Halle [1]1931, Tübingen [4]1972).
Ingarden, R., *Der Streit um die Existenz der Welt*, Bd. II/1 (Tübingen 1965).
Küng, G., „Husserl on pictures and intentional objects", *Review of Metaphysics* Bd. 26 (1973).
Küng, G., „The phenomenological reduction as epoche and as explication", *Monist* Bd. 59 (1975).
Küng, G.,„What do we «see» when we register a fact?", in: R.Haller und W.Grassl (Hrsg.), *Sprache, Logik und Philosophie*, Akten des 4. Internationalen Wittgenstein Symposiums (Wien 1980).
Mally, E., „Gegenstandstheorie und Mathematik", in: Verhandlungen *des 3. Internationalen Kongresses für Philosophie* (Heidelberg 1908).
Mally, E., *Gegenstandstheoretische Grundlagen der Logik und Logistik*, Ergänzungsheft zur *Zeitschrift für Philosophie und philosophische Kritik* Bd. 148 (Leipzig 1912).
Meinong, A., „Über Gegenstandstheorie" (1904), in: *Gesamtausgabe* Bd. 2 (Graz 1971).
Meinong, A., *Über Möglichkeit und Wahrscheinlichkeit* (Leipzig 1915), in: *Gesamtausgabe* Bd. 6 (Graz 1972).
Parsons, T., *Nonexistent Objects* (New Haven 1980).
Smith, D.W. und McIntyre, R., *Husserl and Intentionanlity,* Synthese Library Bd. 154 (Dordrecht 1982).
Smith, D.W., „Is this a dagger I see before me", *Synthese* Bd. 54 (1983).
Tye, M., „The adverbial approach to visual experience" *Philosophical Review* Bd. 93 (1984).
Zalta, E.N., *Abstract Objects*, Synthese Library Bd. 160 (Dordrecht 1983).

* * *

INTENTIONALITY, REPRESENTATION AND THE BRAIN LANGUAGE

Werner Leinfellner
University of Nebraska, Lincoln

Today, cognitive psychology and modern brain physiology show an interesting tendency to go back to Brentano's concept of intentionality as the meaning-giving act for all intelligent beings. Actually, Brentano's concept consists of two theses: the inexistence thesis and the synsemantic thesis. The latter states that an expression LE^i can be given meaning by other expressions $\{LE^j\}$. The first thesis assumes that if we think, imagine, judge, something is f.i., being thought of, and this something has to be an object existing in our mind. F.i. when we judge something is affirmed or denied; when we love something is loved.[1] Consequently, objects which do not exist empirically, such as unicorns, exist solely in our mind and do not refer to any object. A weaker version of the inexistence thesis demands only, f.i. that the reference relation exists as an activity or function of our mind either mentally or psychologically. This activity distinguishes the mental from the physical. Brentano's second (synsemantic) thesis comes pretty close to Wittgenstein's "language in use thesis." In the sentence "Tom is thinking about unicorns" a unicorn is now not any longer an inexistent object nor does it need a referential function, but 'unicorn', LE^i, gets its meaning synsemantically from $\{LE^j\}$ = a description taken from the mythical narration. Interestingly, brain physiology and psychology have recently taken up the quest for "mentally" inexisting representatives of external objects. Here the individual's intention relation which goes from the intelligent brain or the language user to the external world (like in the medieval personal supposition relation)[2] has been contrasted with representation—a mere physical relation—which goes into the opposite direction.

Thus, intentionality is regarded again as a two-fold relation, intention^1 and intention^2. s^1=intention^1 connects inner mental states (BW) or linguistic expressions (LE) with external states of our world (O). We may express this by saying that linguistic intention (s^1), or reference, creates a semantic tripel <O, σ, BW> where σ is the brainphysiological specialization of s^1. The second type of intentionality, intentionality^2, holds between linguistic expressions LE^i, $\{LE^j\}$, f.i., in "bodies are extended". Intentionality^2 is a "word by context" meaning-giving process: we extract the meaning of LE^i from $\{LE^j\}$; f.i., when we say, that a bachelor (LE^1) is an unmarried man $\{LE^2\}$ or bodies are extended, we may render the meaning-giving process by semantemes or semantic tripels: $<LE^1, s^2, \{LE^2\}>$. The question is again, what is the nature of the semantemes, of the reference relation s^1, of the meaning relation s^2 or of σ? Do mental states (BW) exist, and if they exist, where? One simple answer was offered by the linguistic representational view. According to this view LE^1 linguistically represents O and LE^1 stands for the meaning of $\{LE^2\}$. Intentionalistic views coalesce with the Wittgensteinian view, that the subject or the language user always attaches certain invariant linguistic expressions (LE) to the same, invariantly recurring empirical state of affairs or objects (O) in the case of intention^1 (or reference = s^1), i.e., the language user establishes the descriptive meaning of LE. In the case of intention^2, the language user fixes the meaning of LE^i by, f.i., a definite description $\{LE^j\}$, as Russell has proposed.

Again we may ask, what is the nature of, f.i., the relation s^2 in the tripel $<LE^1, s^2, \{LE^2\}>$? Is it a linguistic, a psychological, physiological or even an unempirical mental relation? Dennet 1969[3], Fodor 1975[4] gave a clear answer to this question: they regarded both forms of intentionality (intention^1 and intention^2) as cognitive semantical functions; this, one may regard as a back-to-Brentano movement. Additionally, it was of great advantage for cognitive psychology to work with optimal design models: it led to a better understanding of intelligence;

intelligence is not, as in the traditional view, a capacity of the mind or of its mental functions, but a synergetic optimization of sensations, memories and of thinking by an inner "semantic" evaluation in order to maximize the individuals and the species' survival. The semantic turn in cognitive psychology consists in the following: Whenever we use linguistic descriptions of mental acts or functions, such as "I think that P" or "I remember P", we evaluate represent or encode, store, transmit, or transform meaning (P), since the function of intelligent individuals is: 1) to represent information from the environment and bestow "survival" meaning on them; 2) to carry information and meaning from one individual to the other individual (inter-individual semantical communication, f.i. by language); and 3) to process, store, and make usable and understandable the meaning of information, either in our (brain) memory or in our language, for example, by fixing the meaning of one expression LE^i linguistically through a context or a definition $\{LE^j\}$. But it turned out to be a great hindrance and disadvantage for cognitive psychology, that the two explanations of intentionality, the empirical one (intentionality[1]) and the formalistic one (intentionality[2]), stayed completely separated and could never be brought together. The formalists freely used the elegant formalisms of transformational grammar or the deductive logic of the predicate calculus to explain the establishing and the transfer of meaning; the empirists used, f.i. the reference relation as explanation. Dennett's robot[5] and Searle's chinese learning stories criticized the inability of the formalists to solve practical survival problems with their formalistic methods. The paradigmatical story uncovers three mistakes of the formalistic approach: 1)Intentionality[2] cannot be based on logic, nor can it be explained by transformational grammar alone. 2) Language cannot work without the cooperation of both forms of intentionality. 3) Language without brainphysiological foundation is useless, if the task of intelligence is to maximize individual survival. The gist of Dennett's story is not simply that without empirical meaning we are helpless, as it is according to the story: You are locked into the control center of an alien robot-spaceship, a windowless room, where on two walls you will find all the encoded informations from the outside in flashing light signals you do not understand and on the two other walls there are the buttons which you should press to maneuver the spaceship into security and maximize your own survival. What the story rather tells us is firstly, that in such a situation the decision maker or navigator is completly blocked, not only because he cannot understand the signals from outside, but because he cannot evaluate them. This shows clearly that intention[1] is an inner process where we bestow our survival values on the signals or the mental representations of external input. We will see later that this is exactly the case with the brain-wave patterns which we can now compare with the flashing signals in our control room. Thus, we come now to the second interpretation of our control room as the human brain. The brain represents the incoming sensory inputs in form of brain wave patterns (BW), like the flashing signals in our control room, but we are automatically able to bestow on, or associate with them a certain survival value. That is exactly the difference between frequency patterns like sound-wave patterns picked up by our ear and brain-wave patterns (BW). The latter alone get meaning, in a fundamental primary sense, by the survival values and in a secondary sense by association.

The survival problem of our space-ship navigator could even not be solved, if all the light indicators were labelled in your own empirical language, which has intention[1]-meaning only in your well known biological niche, but not in a completely unknown environment, to which we are not adapted at all. Only if our navigator can step outside and acquire intentional meaning[1] in an ostensive way and can evaluate the alien evironment, with respect to his survival chances, he may be saved.

For such and similar purposes and, more generally, to bridge the gap between intention[1] and intention[2] E. and W. Leinfellner[6] proposed 1978 a two-level semantics, which recursively connects the external empirical meaning, the "operative meaning" or intentionality[1], with the internal or "operational meaning" (intentionality[2]) with the help of semantic tripels and n-tupels.[6] Similar semantic systems have been discussed by Suppes[7], Wilks[8], and by the computer semanticists.

We see the following from this story: 1) if intentions[2] (the mere light signals) or, in the

brainphysiological interpretation, the brain wave patterns (BW) are not somehow connected with the external world by intention[1]—this is certainly not the case for healthy humans—there is no survival chance. 2) Intentions[1] (external) or (internal) intentions[2] alone, if they function separately and independently, do not work (Rorty, Dennett, Searle). Intention[1] and intention[2] of intelligent beings are always relative to their specific biological niche. 3) Only if the two levels of semantical functions, operative or intention[1] and the operational or intention[2], are connected and cooperate, our spaceship can be saved. 4) Intention[1] and intention[2] serve a common cognitive purpose: intelligent problem solving. 5) Only together they maximize survival. 6) Intelligence (contrary to knowledge) consists of effective operative-operational semantical decision-making and problem-solving for the purpose of (species) survival. Thus, operative semantics (semantics[1]) and operational semantics[2] have to cooperate in any intelligent problem-solving; for this purpose, they can be connected on a linguistic level, as Leinfellner W. and E. have shown. Since problem solving optimizes survival, the next question is: how do intention[1] and intention[2] cooperate on the lower brainphysiological level of animalic mammalian intelligence and on the human psychic level without the help of a spoken language? If one could answer these questions, then one could regard intelligence and intentionality as the same, i.e., intentionality in both forms = intelligence, as Dennett has proposed it.

5. Intentionality and Perception

One central problem of cognitive psychology has been the question: How shall we explain the connection between the brainphysiological nervous level of sensory perception and the linguistic level? What is the nature of a brainphysiological representative (BW) when we represent, or recognize, or name an object? Does the representative (BW) "inexist" in our mind or our brain? Richardson (1981)[9] has recently proposed to regard external intentionality (intention[1]) as aboutness or ofness of any intelligent system which has to cope with its evolutionary dependency in its specific environmental niche. In a next step, we want to amalgamate a neurophysiologically based intentionality[1] with a linguistic psychological intentionality[1], following Brentano's ideas.

D1. Intentionality[1] is a general semantic or meaning-establishing, low-level process whose first level consists of neurophysiological functions (σ) which map (= /) environmental input (SW) onto an internal brain language (BL) or: SW/BL. The second level maps (λ) brain-wave patterns BW unto linguistic expressions LE^i. Thus we get: $< BW^i, \lambda, LE^i >$. A brain language consists of dynamic fluctuating states of the brain (or brain-wave patterns = BW, $BW \varepsilon BL$; see below) which serve for 1) communication between the neurons 2) for representing σ (encoding) the sensory input, f.i. in the form of light- or sound-wave patterns (SW), in brain wave patterns (BW), $<SW^i, \sigma, BW^i>$, which constitute the brain-language (BL), 3) to produce survival maximation.

D2. This neurophysiological-psychological activity functions in such a way that the whole process can or may store even that which empirically does no longer exist (Chisholm's contingency condition).

D3. Sensory perception must be such that it represents invariantly recurring objects (O) via invariant light-wave patterns (SW) (informational input) onto invariant brain-wave patterns, i.e.: O/SW/BW, or, in terms of tupels, since O = SW: $<SW^i, \sigma, BW^i>$. Only in case of a cerebral resonance of brain-wave patterns (BW) with linguistic patterns or expressions LE we get the tupel: $<BW^i, \lambda, LE^i>$ and by recursion we get: $<SW^i, \sigma <BW^i, \lambda, LE^i>>$

D4. The purpose of this intelligent process is survival maximation.

D5. In living intelligent mammalian brains, representational states (brain waves, BW) function as first-level representatives of sensory input (SW). Once activated, they continue to exist permanently as dynamic (unconscious) frequency patterns (BW) in the mammalian brain. If mapped into the spoken language ($BW \varepsilon BL$, BL/L), f.i. in the human brain, they become conscious.

D1 guarantees that mental acts are distinguishable from physical events, objects, and stimuli; D5 rules out external deterministic causes or causation from outside in a strict deterministic sense, but permits the partial or statistical causation of a stimulus-response type. Thus, respresentation and intention[1] need an object, but do not need a mystical reference, nor an inexisting symbol s. They rather form a physico-physiological, two-level, dynamic process where sound- or light-wave patterns (SW) are represented by brain-wave patterns (BW), SW/BW, and brain wave patterns by invariantly used linguistic expressions (LE), BW/LE. This explains the tripels <SW, σ, BW> and <SW, λ, BW>, which we have used earlier. It was Dennett[10] and Beaty[11], who compared intentionality[1] with optimal design models, as expressed by D4 as the purpose of intentional perception. Maynard-Smith[12] introduced dynamic optimal design models to explain even the evolution of intelligent behavior (Leinfellner 1982[13], 1984[14]); hence, optimal design models for improving survival can be viewed as theoretical reconstructions of intelligent animal and human behavior which begins with intentional perception.

What does this mean? We will see here, for the first time, that human semantic intentionality[1] is a neurophysiological-psycho-linguistic process whose first level can even function without (linguistic) consciousness. 3) This idea got an astonishing confirmation by the representational theory of visual intentionality by Churchland and Gibson which we will combine here with Basar's and Adey's brain-wave pattern theory (see next chapter). Gibson's "Ecological Approach to Visual Perception"[15] states the following: 1) Observers can perceive (see) objects only, if they possess a neurophysiological functioning, screening representational device with a sensory input function (σ) which transforms and structures sensations—which are light- or sound-frequency patterns (SW)—into brain-wave patterns (BW). 2) Dynamic internal states or brain-wave patterns (BW) do not exist alone, but in triplets: $<SW^i, \sigma, BW^i>$, where SW^i is an invariantly recurring, f.i. sound-wave pattern, a melody, σ transfers the sound pattern from the ear to the brain and BW^i is its invariant resonance in the form of an electromagnetic brain-wave pattern. This view is an active alternative to Hume's passive theory that all what we are sensing are our bodily impressions in the ear. Gibson's ecological, perceptual intentionality demonstrates clearly that in the act of perception outside and inside (the brain) are one interconnected holistic dynamic resonating field, as symbolized by the tripel $<SW^i, \sigma, BW^i>$. Perceptual intentionality means that our brain selects certain light-wave patterns and does not represent all light beams from the outside. Thus we see perhaps only 10% of our world, because there is the inner, inbuilt evaluation (screening) of biologically unimportant visual inputs from outside. Visual perceptions are directed by genetically stored "evaluating filters" which are anchored in our genes. They bestow a primary meaning on them and function in the following way: What is not necessary for survival, we do not see and represent in short, has no meaning for us. Eyes are not at all pricked by light rays and do not react always in an automatical, deterministic, causal way; they are mobile scanners; they run over the array just like hands run over the material to feel and evaluate only that what is meaningful for us, i.e. for survival support. (Gibson 1968)[15]. Neutral experiences do not impress us and get no primary meaning. Therefore, light and light representation are completely different. Light-wave pattern representation is intentionally manipulated and directed by the "seeing" individual brain and is uniquely specific for our niche and scenario in which we had to survive since millenias. Thus we are able to explain for the first time, why the intention[1] relation goes from our brain to the objects, and why representation works in the opposite direction. But perception functions only in cases of "intentionally admitted" and regulated references, admitted by the "brain" to get all necessary lifesaving information from the external world. Therefore, it is the first step of a problem-solving and decision process which maximizes our survival and wellbeing. Thus the representational theory has changed from a passive view of reception to an active view of intentionality[1] which depends on the subject's survival or survival instinct. This intentional model of perception vindicates Brentano's psychological view of intention and could even explain neurophysiologically the lower-level function of intentionality[1].

Gibson's model of perception and similar other intentional perception theories can explain

the following open questions of Brentano's intentionality: 1) the strategic role of the intentional for all mental life which Brentano, as Spiegelberg[16] says, had discovered, but could not explain and explore; 2) The new perception theories are based on the neurophysiological nature of intentionality[1] and stress their general semantic task. 3) They explain the direction of intention[1] which originates from the brain's and from the sense organs' scanning activity and which actively selects and modifies sensory informations from outside. 4) The individual is no longer a passive recipient, since Gibson's model shatters the deterministic stimulus-response theories of perception. 5) Given intentional perception, we need the intermediary function of the brain language (BL) which will serve as the missing link between sensory input (SW) and our spoken language (L), according to the representational two-level mapping: SW/BW and BW/LE. or: $\langle SW^i, \sigma, BW^i \rangle$ and $\langle BW^i, \lambda. LE^i \rangle$.

7. The Internal Language of the Brain: The Brain Language as the Missing Link

New results about the communication of the neurons in human brains have rejected the traditional idea that the brain functions just as a computer or a switching network, which consists of 25 billions of individual processing and firing cells, the neurons, which communicate with each other in a linear and pairwise fashion through electrical impulses. Since any language is a system of signs or signals for the purpose of communication, we will call their exchange of meaningful information the single-cell communicative language. According to the single-cell or feature extract model of the brain, neurons transmit electrochemical impulses from one to another via long fibers, the axions. The impulses travel along the axions until they reach the junctions called synapses which connect the axions with other cells' axions. Chemical transmitters, released from the axion terminals, cross the synapses and excite or inhibit the connected neuron. If the excitation from several synaptic inputs exceeds a certain threshold, the target neuron fires or "switches on" by generating an impulse of its own, or "switches off".

Thus, the brain was inevitably compared to a digital computer, although scientists pointed out that this cannot be the case: f.i. millions of neurons can not any longer digitally encode all the other neuronal impulses, since studies of sensory neurons showed that the single-cell communication consists simply in the increase or decrease of firing intensity which does not work digitally, but pulse code analog. More rapid firing simply represents more intense sensation. Since the impulse frequency varies continuously, sensor neurons rather use a pulse code analog, not a digital form of information transfer.

Fascinated by the computer analog model of the brain, scientists studied the internal language of single neurons, but never of cooperating cells in groups. They could explain the linear language of neurons by gradually beginning with the pairwise neuron communication and then continuing with neuron circuits, but they finally failed to explain the communication between all ganglions of the brain as it has been observed in special stereo-electroencephalograms and in evoked potentials of the brain, the EEG-EPrograms (see Basar, 1980 p. 22–29)[20]. The spontaneous electrical activity of the brain (EEG) can be influenced by direct or peripheal stimulation (electrical, optical, acoustical, linguistical, etc.) and yields the evoked brain-wave patterns (BW). Given that a group of scientists, E. R. John[17] (New York University Medical Center), Lawrence and Adey[18] challenged the single-cell communication or the computer model of our brain, because it assumed that at the top level of the neuron chains there had to exist highly specified single neurons, f.i. with the ability to recognize a person's body as a whole or as a gestalt. But how could one single cell synthesize millions of sensations of different types into one whole? Holistic systems or gestalts cannot be represented in specific single neurons, on the contrary, recent operations on the brain which removed large sections of the brain, where these highly specified neurons were assumed to be located, proved that the individual's memories were merely slightly impaired and not destroyed. This proved that memories must be distributed throughout the brain, that they had to be ubiquitous within the brain! It is clear that the computer-analog model could not explain this. It

could only explain that primitive forms of communication in the single-cell language are similar to the machine language of computers. For example, if switches are shut off and on, then the basic negation function is realized in the single cell communication language just by the analogy of firing and nonfiring of a neuron. It could explain how "and" or the intersection communication of signals are realized in serial communication circuits of neurons and how "or", the union communication, worked in simple neuron networks in parallel circuits of neuron chains. "All" statements or even fuzzy communication are realized by long series of parallel neurons. Thus it was possible to explain how simple syntactico-semantic operation like "non", "and", and "or" combination of signals, work in the single cell communication, namely, in invariant switching structures. But it could not explain communication between neurons which do not belong to the same linear chains, or between large groups, or between all neurons; further, it could not explain, how neuron networks represent similar or statistically recurring events, properties, or sensations, or how our brain could represent, associate, and compare similar sensory inputs, such as gestalt features, patterns, etc. Since the early 1970's, neuroscientists found out that there is a far more complex form of functioning intercommunication between our brain neurons, the cooperative one. According to this view, neurons have the ability to cooperate and communicate in large groups. They produce and communicate with the help of complex electromagnetic fields which pervade the brain and serve as a "collective" internal brain language between the neuronal body of thousands or millions of synchronized and cooperative neurons and at the same time could be understood by single neurons.

The cooperative model, outlined by John and Adey and E. Basar[19, 20], emphasizes the cooperative network communication rather than the individual linear communication. It grew out of the study of evoked electroencephalograms (EEG-EPrograms) which show clearly the measurable, slow electromagnetic wave patterns which pervade the brain. This special EEG can be recorded by electrodes and observed on screens. Research in the 1960's demonstrated that these special EEG's arise not from the random summation of individual cells, nor from linear neuron chains, but from a cooperative organized brain language of large groups of neurons with other groups. Each of the slow electromagnetic communication fields, produced by the individual neurons and by the different groups, are characterized by a different, or similar (identical), wave pattern BW. We will call both communication systems, the single-cell and the cooperative one, the internal brain language, BL, a concept which has been introduced by Jaynes to explain the internal communication between the two hemispheres of the mammalian brain.

This cooperation yields the formation of encoded, relatively invariantly recurring communication patterns of electromagnetic waves BW, whose message can be "understood" by the single neurons which belong to the cooperative group (BW^1) or, via resonance, to a similar group (BW^j). It represents in the first-level, internal brain language, BL, the incoming sensations. Invariantly swinging sound- or light-wave patterns (SW), coming from objects in our environment (= sensory informations), are represented (σ) encoded, and stored in the corresponding brain-wave patterns (BW), or form the tripel <SW, σ, BW>. They originate from specific areas of our brain and spread over the whole brain. The wave patterns show, on the one side, significant structural similarities with the second level of representation, the linguistic expressions LE. The latter, too are nothing else than similar, invariantly swinging sound-wave patterns, which are activated by the Wernicke center of our brain and produced by the vocal apparatus whose muscles are regulated and controlled by the Wernicke center. Actually, linguistic expressions LE in the tripel <BW, λ, LE> occur 1/300th second after the brain waves: Brain-wave patterns (BW) explain, on the other side, how it is possible that the most important external stimuli—which again are nothing else than invariantly recurring external sound- or light-wave patterns (SW)—can be represented by invariantly swinging brain-wave patterns BW: SW/BW. Thus, the picture theories of representation and similar older explanations of sensory representations, as well as the whole sense data theory become more or less obsolete.

Adey and Soviet scientists demonstrated recently 1) that specific complex electroencephalogram (EEG) patterns can be correlated with what we call second-level specific cognitive processes both in animals and in human brains (= ABW). They could explain, how certain invariantly given sensations (WS) are stored in dynamically swinging, invariantly recurring simple (BW) or abstract wave patterns (ABW). Growing evidence supports the idea that the first level, the invariantly swinging electromagnetic brain-communication patterns (BW), represent and store (isomorphically) certain invariantly, i.e. in the same shape, recurring sensory frequency patterns. Moreover, they could prove that the lower (BW) and the higher brain-wave patterns (ABW) are invariantly "translated" into linguistic expressions LE of our language L. Thus we get: BW^i/LE^i, ABW^j/LE or $<BW^i, \lambda, LE^i>$, $<ABW^j, \lambda, LE^j>$. According to our brain-wave hypothesis, the overall pattern of electromagnetic brain-waves (BW), representing sensory wave patterns (SW) or stimuli, firstly represent them invariantly; secondly, the fact that brain-waves (BW^i) resonate (or do not resonate) with other invariantly given similar wave patterns BW^j explains, how similarities (and dissimilarities) between sensations are distinguished in the brain. Similarities lead to resonances just like tuning forks resonate with specific frequencies. Abstract brain-waves (ABW) are superimpositions of brain-waves (BW), the represent millions of stimuli; if the latter are sound- or light-wave patterns (SW) does not matter. Moreover, communicative brain-wave patterns can activate or "recall" similar swinging patterns by resonance (=) even if they are already stored away in the chemical dynamic structure of other groups of neurons—and increase their intensity. But $BW^i = BW^j$ resonance is a typical non-linear oscillation and activation (under random) of similar stored and swinging wave patterns. Activation by resonance brings them back from their low energetic level to the threshold value at which they enter our consiousness (we did not discuss here the direct chemical senses (smell) which function through chemical interaction and transfer of chemical stimuli). Thus the brain-wave pattern (BW^i) (of electromagnetic waves) is the representation of an incoming sensory sound stimulus (sound-wave pattern SW^i). F.i., hearing the same song (an invariantly occurring sound-wave pattern) is represented by an invariantly swinging brain-wave pattern BW^i (SW^i/BW^i), and this brain-wave pattern invariantly retains its stored characteristic pattern and frequency. The pattern BW^i may trigger BW^j, and BW^j may resonate with a linguistic brain wave pattern LE^j, and the words of the melody BW^i will come now into our consciousness simply by the increase of BW^i's intensity. We have to assume the following two-level linguistic hypotheses: 1) synchronicity hypothesis: the internal brain language (BL) can be represented by, or translated into the language L: BL/L. In this case brain-wave patterns (BW) resonate and stimulate our vocal muscles and activate speech patterns, vice-versa. 2) Diachronicity thesis or Jaynes' thesis: our acoustic language originated historically or evolutionarily, step by step from the brain language. 3) Consciousness hypothesis: we may go so far as to assume that, if the stream of invariantly swinging brain-wave patterns (BW, ABW) has enough associative (communicative) power to resonate and activate similar "linguistic" wave patterns (LE), then they will become conscious to us. Thus the stream of consciousness is born and may show up as a sequence of spoken and heard linguistic expressions {(LE)}. (Jaynes). The role of the neurons essentially boils down to creating and producing the energy for, and to modulate the intensity of the already stored as well as of the new emerging brain-wave patterns. 4) The invariantly given specific brain wave patterns and the specific differences between the invariantly swinging innumerous different patterns can be seen on EEG screens and can be heard in the corresponding combinations of the linguistic expressions (LE), since the specific brain-wave patterns occur in the brain always 1/300th of a second before the corresponding words are spoken. In recent experiments, the specific brain-wave pattern for "switch on the light", appearing in an EEG has been amplified and used to directly switch on a light in a room with the person remaining silent! Thus the brain-wave "switch on the light" exists, and it can activate an electric switch without the use of the language. 5) Invariantly swinging brain-wave patterns represent the semantical contents of external, invariantly swinging sound- or light-wave patterns interpersonally, i.e., they are the same in different persons as experiments have convincingly shown.

We have to assume, that all newly emerging and representing communicative invariant patterns (BW) change the chemical and electric structure of the neuronal group, i.e., of the neurons, and thus have fixed memories in a new, dynamically and invariantly swinging brain-wave pattern. Wave patterns which are not reinforced by resonance will decrease their energy level slowly to a threshold value, where they will become subconscious or where we forget them. But they can be brought back if stimulated by new associative, enforcing resonance. That is exactly, how the dynamic retrieving of memories by resonances works.

Any of these retrieved amplified "memories", or, brainphysiologically speaking, any of these activated swinging neuronal wave patterns can be converted back via impulses along the axions, which again can be transmitted to the appropriate muscle cells to yield a motoric response in the organism (body): I will switch on my record player to listen again to the song, or begin to speak.

We have to stress again the fact, that, if all the items of information in the brain are encoded in the frequency spectra of modulated brain-waves, and if the rest of the brain cells and groups understand this message which is encoded in the brain-wave patterns, then the language of the brain 1) is ubiquitous, i.e., is everywhere in the brain and can be understood by all neurons; 2) both languages, the brain language (BL) and the spoken language (L), have a similar and invariant frequency structure. 3) Their semantic content has sentence like structure, but we have to assume that in the brain language "noun-parts" and "verb-parts" have equal weight. Therefore, the brain language (BL) may not differ functionally too much from a spoken language (L) which is encoded in the changing, but invariantly used, spectra of sound waves, in the words and their combinations, and is understood—just as the brain language by all neurons—by all language users. The salient point to understand the functioning of the brain-wave language is that the abstract brain-waves (ABW) and the simple brain-waves (BW) represent and encode the semantic content of what we call "sensations", "memories", "thoughts". Recent experiments have shown that it is solely their meaning for us which is encoded and represented. When Warren Brown recorded brain-waves of subjects, who listened to identically sounding words (homonyms) which are used both as nouns and verbs, such as "rows" in "he rows" and "rose" in "the rose" the EEG responses, i.e. the brain-wave patterns, were significantly different! Thus the basic assumption of the brain language is a semantic version of the inexistence thesis of Brentano. According to this thesis, the healthy human brain maintains a life-long, lasting activity of brain-wave patterns which represent, compare, combine, and store cerebral meaning, generated by the cooperative, dynamic activity of thousands or millions of neurons themselves. The encoded cerebral semantic items of information in turn influence the activity of all neuron cells. Therefore, this model uses a new dynamic form of inexistence, not of mental objects, but of cerebral meaning invariances, firstly in the form of operative or empirical meaning, when certain invariantly swinging brain waves represent or stand for invariant sensory stimuli (in the sense of SW/BW representation), and secondly, in the form of a cerebral operational semantics or prelinguistic meaning-establishing and storing, when certain brain-waves, f.i. BW^i, are identical or similar with others, BW^j, i.e. $BW^i = BW^j$. The most astonishing result of this new brain-language theory is certainly that the brain language is based completely on, and functions exclusively in the form of cerebral semantic relations which order and fix the cerebral operative (empirical) and operational meaning. We do regard here the brain only in the context of intelligence (not of knowledge or science). Only for intelligent behavior, the establishing and fixing of meaning is the first necessary step for the final goal, the maximation of survival and of welfare. (see Leinfellner, W., 1984, note 14, p. 264)

6. Episodic and Semantic Memory and the Parallelism between Cerebral and Linguistic Semantics.

According to Tulving's[21] propositional memory theory, there is one totally wrong memory theory: the static-warehouse theory, where the owner can find what he wants to find (Gregor

von Feinagle's concept, 1807). Thus the warehouse concept is exactly the static theory which assumes inexistent mental objects and which Brentano gave up in his later years. We want to combine now the theory of the dynamic brain language with Tulving's semantic theory. According to the latter, memory is stored in the episodic memory which is exactly an operative semantics (semantics^1) and in the semantic memory, which is strictly functioning as an operational semantics (semantics^2). We stress the following: 1) There is no such thing as "the static memory", but different forms of inexistent, permanently activated neurophysiological brain states (processes) or brain-wave patterns, spread over the whole brain, whose semantic content may easily show up in our language. 2) Retrieval of information is not finding a stored good, an inexisting mental object; it is the activation of brain-waves of low intensity which, if activated, appear in our consciousness as linguistic expressions. This emergence is not always successful; otherwise the rememberer would know what he has forgotten. 3) Memory is not determined by past memory contents; memory is, rather, a two-levelled semantic activation process with the goal to "translate" (λ) the dynamically swinging brain-wave patterns (BE^i) into linguistic expressions (LE^i). Thus we get the tripel $<BW^i, \lambda\ LE^i>$ or BW^i/LE^i.

But we have to distinguish the procedural from the propositional linguistic memory. Kolers 1945[22] and Scheffler 1965[23], distinguish in a similar way between procedural memory and a semantic or substantial memory, or between skills and intelligent knowledge. Skillful procedures are neither true nor false, since we just execute automatically a brain-wave pattern; but we find and reveal words in our linguistic memory. Executing a skill means to follow a brain-wave pattern instinctively. Therefore we do not necessarily need linguistic instructions when we perform skills; we do it automatically, instinctively, and unconsciously, like adding or multiplying numbers. Propositional or linguistic memory can be subdivided, according to Tulving, into episodic and semantic memory. The first is practically a generalized intention^1 or semantics^1, the second is an intention^2 or a semantics^2. Both convert the cerebral semantics of the brain-wave language into linguistic semantics. Episodic memory is encoding and retrieval of personal experiences and happenings; but the semantic memory encodes and retrieves the invariant, semantic rules, the cerebral semantic networks of meaning for the purpose of how to use and understand meaning, independent of a person's identity and past; i.e., semantic memory activates via the cerebral the linguistic functions of the semantics^2. For we recollect a car in which we were riding, but recall an identical meaning or a semantic rule. Episodic memory is a time-creating, dynamic system which retrieves and stores brain-waves (BW) and the corresponding linguistic expressions (LE) by ordering them in a linear temporal and a spatial, locational frame or raster. Thus, the representatives, the brain-wave patterns (BW), impose their temporal and spatial order in an "empirically apriori" sense on sensations. The semantics^2 (cerebral and linguistic memory) stores meaning networks and rules necessary for the meaningful use of both languages. They resemble a "mental" Webster's Dictionary, a thesaurus, an organized semantic knowledge which a person possesses about the meaning (intention^2) or words and verbal symbols: how their meaning is described or defined by other words or contexts and how one may connect them via the episodic memory and the brain language to the referents (SW). It activates semantic relations among words, sentences, "switches on" rules, formulas, algorithms for the understanding and manipulation of brain waves, insofar they show up in spoken words or in their combinations.

There are no external input signals registered and necessary, once semantic memory is stored; this proves its mere internal, "linguistically apriori" semantic functions. But for the cognitive linguistic expressions quasi-referents are stored and obtainable from the episodes of the episodic memory. Retrieval from the semantic memory will leave its semantic networks (rules) unchanged. The two memories are interdependent, but nevertheless they mutually support each other in the sense of a linguistic, cognitive recursivity of internal intentionality^2 and external intentionality^1, i.e., $SW^i/BW^i/LE^i$, LE^i/LE^j. To make it clear: the linguistic empirical information stored in the episodic memory, f.i., is always about the past temporal co-occurrence of two words, whereas, f.i., the identity of 'body' and 'extension' is stored in the seman-

tic memory as experienced (invariant) synonymity: "all bodies are extended". Therefore, what logicians call "analytically true" turns out to be "having the same meaning". The semantic networks can be reconstructed in recursive or in computer semantics (Leinfellner, E.)[24] which makes it possible for computers to utilize their knowledge in a semantic, and not logical form as well as to understand our language. Since the semantic, cerebral memory is phylogenetically and ontogenetically earlier than the linguistic memory, this can only be seen as a proof that the cerebral "silent" semantic memory existed already before the spoken language, i.e. as the cerebral semantics[2] for the non-spoken, communicative brain language (BL). The items of information stored in the episodic linguistic, semantic[1] system are the personal, individual experiences, while the information in the semantic[2] memory regulates their semantic use. The first is limited in its inferential capabilities, but the semantic memory possesses a rich inferential and deductive, exclusively semantic capability, because of the tightly organized and recorded semantical interdependencies and interrelations, which build up the semantic networks of the brain language and the spoken language (See Leinfellner, E.[24]). But one should never forget that the semantic memory cannot be based on logic, since logic develops ontogenetically and phylogenetically later than semantics. Therefore, logic is the offspring of a brain-physiologically and psychologically based semantics, a thesis already used in Brentano's psychological foundation of logic and logical deduction. F.i., in deductions the conclusions are always semantically included in the premisses, i.e., brain-wave patterns of the premisses have to be composites or superimpositions of the brain-wave patterns of the conclusions. In philosophy, it was Wittgenstein who long ago took the truth tables for "and", "or", and the negation from language use, i.e., from their semantic use and functions in our spoken language. Consequently, the foundation of our linguistic semantics on the cerebral semantic networks and functions of the brain language is the next step in the physiological-psychological foundation of our spoken everyday language (not of the scientific language) on the silent, but semantical, brain language.

Only episodic memory possesses a subjective veridicality which is actually a subjective memory evaluation, a truth testing that invariant names (LE^i) and invariantly used verbs (LE^j) refer to the same invariantly reoccurring and concurring events, SW^i, SW^j. Of course, the rigid and invariant empirical references of our spoken language, such as naming, develop historically, when the spoken language evolved historically, by a step by step activation from the brain language (Jaynes' hypothesis). Thus the empirical truth is nothing else than Brentano's evidence experience, when we retrieve and use the same (invariant) linguistic expression LE^i for the same invariantly (recurring) sensory experience (SW^i), only 1/300th of a second later, than our brain has already attached an invariant brain-wave pattern (BW^i) to the same invariant sensory experience (SW^i). Thus we get the recursion of the tripels $\langle SW^i, \sigma, BW^i\rangle$ and $\langle BW^1, \sigma, LF^i\rangle$ to $\langle SW^i, \sigma, \langle BW^i, \lambda, LE^i\rangle\rangle$ or, simply, in our symbolic translatory form: $SW^i/BW^i/LE$.

Finally, it is very well known that platonistic logicians assume that logical truths and tautologies are apriori given as timeless truths, independent of all sensory or empirical experience. But they have no explanation, why whatever is apriori must possess universal and necessary validity in all possible worlds. The brainphysiologically based semantics explains apriorism quite differently, namely as genetically inherited, statistical invariance of our cerebral and linguistic semantic functions and networks, which have gradually developed, during the evolution of the intelligent human brain, our semantic universe of both languages. Since, according to Tulving's memory theory, only our episodic memory is temporally ordered, while the semantic memory is timeless—like logic—it does not need any longer temporal connections to the knower's personal empirical history, but it could be so connected, if necessary. Instead it uses the evolved, established and inherited invariant semantic functions and forms, the universe of "discourse" of our brain language on the first level, and/or the synchronized semantic universe of our language at the second level. It does not refer to logical, apriori given, and moreover unempirical possible worlds. The timeless and intergenerational, genetically inherited, invariant semantic functions and networks of the brain language are guaranteed and

maintained by the genetic inheritance from generation to generation. The semantics of our everyday language continues to function correctly, if we use it and if it is written down in our cultural tradition or in our textbooks. Therefore, the invariance of references, of established meanings, of meaning inclusion, of invariant similarities (synonymity) and of meaning are the physiological and psychological preconditions of any logical laws and, of course, of "analytic truth". Thus the possible worlds turn out to be the semantical universe of both languages, the brain language and the spoken language. There are no objections against the most abstract or obstruse bits of linguistic semantic memory knowledge. It may be a creative combination of fantastic or abstract wave patterns, or it may once have stood in a direct connection to our external environment in which we and unicornlike, and dragonlike creatures coexisted and lived. Today only the meanings of the "unicorns" and "dragons" live on in our cultural, mythical tradition and their once formed brain-wave patterns roam about in our semantic memory, but fortunately not any longer in our episodic memory.

ENDNOTES

[1] Brentano, F., *Psychologie vom empirischen Standpunkt* (Leipzig 1874), Vol. I, Book II, Chap. 1.
[2] Leinfellner, E., "Intentions as *intentiones animae* in Ockham and as Immanent Mental 'Objects'", this volume.
[3] Dennett, D. C., "Current Issues in the Philosophy of Mind", *American Philosophical Quarterly* Vol. 15 (1978), pp. 249–261, especially p. 257f.
[4] Fodor, J., *The Language of Thought* (New York 1975).
[5] Dennett (1978), p. 258.
[6] Leinfellner, W., and Leinfellner, E., *Ontologie, Systemtheorie und Semantik* (Berlin 1978), pp. 216–359.
[7] Suppes, P., "Procedural Semantics", in R. Haller and W. Grassl (eds.), *Language, Logic and Philosophy* (Wien 1980), pp. 27–36.
[8] Wilks, Y. *Grammar, Meaning and the Machine Analysis of Language* (London 1972).
[9] Richardson, R. C. "Internal Representation: Prologue to a Theory of Intentionality", *Philosophical Topics* Vol. 12 (1981), pp. 171–211.
[10] Dennett (1978).
[11] Beatty, J., "Optimal Design Models and the Strategy of Model Building in Evolutionary Biology", *Philosophy of Science* Vol. 47, pp. 532–561.
[12] Maynard-Smith, J., *Evolution and the Theory of Games* (Cambridge 1983).
[13] Leinfellner, W., "Das Konzept der Kausalität und der Spiele in der Evolutionstheorie", in K. Lorenz and F. Wuketits (eds.), *Die Evolution des Denkens* (München 1983), pp. 215–261; W. Leinfellner, "Evolution of Intelligence", in P. Weingartner and J. Czermak (eds.), *Epistemology and Philosophy of Science* (Wien 1983), pp. 161–168.
[14] Leinfellner, W., "Theory of Games and Evolution of Intelligence", in F. Wuketits (ed.), *Concepts and Approaches in Evolutionary Epistemology* (Boston 1984), pp. 223–276.
[15] Gibson, J.J., *The Senses Considered as Perceptual Systems* (London 1968), p. 251.
[16] Spiegelberg, H., "'Intention' and 'Intentionality' in the Scholastics, Brentano and Husserl", in L. L. McAlister (ed.), *The Philosophy of Brentano* (Atlantic Highlands, N.J. 1977), pp. 108–127, especially p. 108.
[17] John, E. R., "Multipotentiality: A Statistical Theory of Brain Function", in R. J. Davidson and J. M. Davidson (eds.), *The Psychobiology of Consciousness* (New York 1980).
[18] Lawrence A. F., and Adey, W. R., "Nonlinear Wave Mechanisms in Interactions between Excitable Tissue and Electromagnetic Fields", *Neurological Research* Vol. 4 (1982).
[19] Basar, E., et al. (eds.), *Synergetics of Brain* (Berlin 1983).
[20] Basar, E., *EEG-Brain Dynamics* (Amsterdam 1980).
[21] Tulving, E., *Elements of Episodic Memory* (Oxford 1983).
[22] Kolers, P. A., in L. S. Cermak and F. I. M. Craik (eds.), *Levels of Processing in Human Memory* (Hillsdale, N. J. 1979).
[23] Scheffler, I., *Conditions of Knowledge* (Glenview, III. 1965).

[24] Leinfellner, E., Steinacker, I., and Trost, H., "Anaphoric and Kataphoric Relations, Textual Cohesion, and Reference Resolution", *Computers and Artificial Intelligence* (in press); R. Trappl, E. Leinfellner, I. Steinacker, and H. Trost, "Ontology and Semantics in the Computer", *Language and Ontology: Proceedings of the 6th Wittgenstein Symposium* (Wien 1981).

I have to thank Prof. Basar, Institute for Psychology, Lübeck, for his helpful comments, E. Leinfellner for her semantic contributions to the research project, and finally the Research Council and Dean Meisels, University of Nebraska, Lincoln, USA, for the financial support of this research project.

* * *

WHO, SEARLE OR HUSSERL?

A. Z. Bar-On
Hebrew University of Jerusalem

The question dealt with in this paper, broader than that of the title, can be phrased as: who is to guide us in analysing the concept of intentionality? I shall try to answer it by tracing a path of thought leading from Brentano, through Husserl, to Searle, which is one of a few possible paths that can be shown to bring us from Brentano, as the source, to the theories of our days.

I

It is presumably agreed that it was Brentano who proposed the initial thesis of intentionality in modern times. He did it in two short adjacent passages in the second book of his *Psychologie vom empirischen Standpunkt* (quoted from the translation of C. Rancurello, D. B. Terrell, and Linda L. McAlister)[1]:

> Every mental phenomenon is characterized by what the Scholastics of the Middle Ages called the intentional (or mental) inexistence of an object, and what we might call, though not wholly unambiguously, reference to a content, direction toward an object (which is not to be understood here as meaning something real[2]), or immanent objectivity. Every mental phenomenon includes something as object within itself, although they do not always do so in the same way. In presentation something is presented, in judgement something is affirmed or denied, in love loved, in hate hated, in desire desired, etc.
>
> This intentional in-existence is characteristic exclusively of mental phenomena. No physical phenomenon exhibits anything like it. We can therefore define mental phenomena by saying that they are those phenomena which contain an object intentionally in themselves.

Brentano attached to this text two footnotes which are quite significant for its proper understanding.[3] In the first of them, the one attached to the place where the Scholastics of the Middle Ages are mentioned, Brentano reminds us that they also used the expression 'to exist as an object in something', but in his view we should refrain from employing it, because of the difference between their and our usage of the term 'objectively'; ours suggesting *real existence*, while theirs intimating just the opposite.

In the second footnote, attached to the end of the first passage, Brentano claims that the conception of inexistence can already be found in Aristotle, as much as in the writings of several later thinkers, like Philo, the Neo-Platonists, St. Augustine, Thomas Aquinas.

As to the text itself, we should, I think, distinguish in it the main definitional thesis from the elucidations and/or extensions added to it. In the thesis the single characteristic of mental phenomena, intentionality, is referred to by no less than four expressions which are supposed to be co-extensive: (1) intentional inexistence; (2) reference to a content; (3) direction toward an object; and (4) immanent objectivity. These expressions are supposed to describe the same property or relation, but they cannot, it seems to me, be considered synonymous. Each of them has some secondary connotations and only their joint application will prevent misunderstandings.[4]

As it is commonly acknowledged, the most important of these parallel expressions is the first one, connecting the conception of intentionality with certain traditions of the Middle Ages. Brentano's "mediaeval connection" need not be told here in its entirety, but I cannot

omit it altogether. In the second footnote to the paragraph quoted above Brentano says: "Aristotle himself spoke of this mental in-existence. In his books on the soul he says that the sensed object as such is in the sensing subject; that the sense contains the sensed object without its matter; that the object which is thought is in the thinking intellect".[5]

What is meant here, as I read the text, is Aristotle's epistemological conception, according to which the knowing subject, whether through his sense-experience or by his understanding, takes in the *form* of what is being apprehended, without its matter. Hence, this conception assumes a threefold distinction applicable to the form: (a) the form *per se*; (b) the form *as in the thing*, as *its* form or essence; and (c) that very form as in the mind of the subject who apprehends the thing of which it is the form.

Now, Avicenna, who showed a special interest in the *third* sense of the form, had it identified with what in Arabic is called *máana*, i.e. meaning or concept. He distinguished, however, between the primary *máana*, such as 'man', 'ox', or 'donkey', that the subject applies to the man, the ox, or the donkey, which exist outside of his soul, and the secondary *máana*, such as 'species' or 'genus', also meanings or concepts in the soul, but such that the subject applies them not to things existing outside of the soul, but to the primary *máanas*, such as 'man', 'ox', or 'donkey'.[6]

When Avicenna's writings were translated into Latin, the term *máana* was replaced by *intentio* and the distinction between the primary and secondary *intentiones* were elaborated upon, giving thus birth to the adjective 'intentional'. What seems to be uncontroversial about the mediaeval usage of this term is that it was meant to signify a mode of existence different from the so called 'real' existence, i.e. existence in the mind as distinguished from 'external existence'. There is, however, quite an amount of disagreement as to the question whether it also signified a reference or directedness to an object.[7]

One has to take sides in this disagreement in order to answer the question of whether Brentano remained faithful to the mediaeval usage of the term 'intentional'. What, as it seems to me, cannot be doubted, is that *both* of the above mentioned meanings are there in Brentano's concept of intentionality, as characterized in the quoted passage. It can be said that Brentano appears to have combined in his conception of intentionality two distinctions: one between intra-mental and extra-mental existence, the other between directedness and the lack of directedness to an object. This explains, as I see it, the need for additional expressions referring to the characteristic in question. By means of the second phrase, 'reference to a content', Brentano reminds us of the first distinction, 'content' pointing to the exclusively intramental. The third phrase, 'direction toward an object', reminds us of the second distinction; not simply intra-mental existence is considered, but the reference to an object whose existence is intramental. Finally, in the fourth phrase, 'immanent objectivity', both distinctions are intertwined in a single term.

Let us consider now the three elucidations which Brentano had added to his main thesis. The *first* one, in the opening sentence of the first passage of our text, alludes to what we have just said: the alternative expressions are ambiguous and are therefore to be applied jointly, each one contributing its own shade of meaning to the correct understanding of the thesis.

The *second* elucidation can be found in the brackets after the phrase 'direction toward an object'. It is there to tell us how to construe the key-concept 'object' as a component of Brentano's thesis. The object need not be real. Reality is not a necessary condition for something to be intentionally included in the mental phenomenon as an object. Indeed, at this point Brentano does not tell us in what sense the term 'real' and 'reality' are used here. It may, however, be assumed that what is meant here is reality as the state of existing independently of being perceived or experienced in general. This is certainly the sense suggested by the footnote mentioned above. Anyhow, the point is controversial. It attracts the attention of those who are interested in Brentano's thesis more perhaps than any other of its aspects. It is, however, less prominent in our framework of discussion.

The *third* elucidation is what matters to us most. It consists of two parts: (a) what follows the word 'although' in the sentence "Every mental phenomenon includes something as object,

although they do not always do so in the same way; and (b) the examples of mental phenomena quoted by Brentano. These were, no doubt, chosen with care; the collection exemplifies the basic classification that Brentano seems to have adopted for his inquiry. According to it all mental phenomena divide into presentations (*Vorstellungen*), e.g. hearing a sound, seeing a coloured object, feeling warm or cold, imagining a mythological figure, on the one hand, and such mental phenomena what have a presentation as their basic component, on the other. This second class divides in its turn into judgements and phenomena of feeling and desire or will.

II

This third elucidation to Brentano's thesis may be seen as leading us to Husserl's conception. Husserl's theory of intentionality is indeed a thick web of arguments and analysis spreading over his whole philosophical system, its warp the diachronic and weft the synchronic order. We do not intend to deal here with all, or even with most of its facets; our discussion will be restricted to only a few aspects which seem to be particularly relevant to the framework of the present analysis.

Let us begin with a remark on terminology. Instead of Brentano's 'mental phenomenon' of which intentionality is supposed to be a definitional characteristic, Husserl uses the term 'experience' (*Erlebnis*), adding that not each and every experience is intentional. And so, whatever it will prove to be, intentionality cannot be considered a definitional feature of mentality. A class of examples of non-intentional experiences is given by a certain kind of *Gefühle*.

Husserl's *Logische Untersuchungen* read:

> Many experiences commonly classed as 'feelings' have an undeniable, real relation to something objective. This is the case, e.g., when we are pleased by a melody, displeased at a shrill blast etc. etc. It seems obvious, in general, that every joy or sorrow, that is joy or sorrow *about* something we think of, is a directed act. Instead of joy we can speak of pleased delight in something, instead of sorrow we can speak of displeased or painful dislike of it, aversion from it etc. etc.[8]

Note that in Brentano's view in each of these cases there would be two intentions, one of them based on the other. The underlying intention would give us the object presented, while the overlying intention would give us the object felt. The former can be separated from the latter, but not *vice versa.*

Up to this point Husserl goes along with Brentano. He thinks, however, that there are also feelings which are essentially non-intentional. He writes:

> We may now ask more generally whether, in addition to the intentional varieties of feeling, there are not other *non*-intentional species. It may seem . . . that an obvious 'Yes' is the right answer. In the wide field of so-called sensory feelings no intentional characters can be found. The sensible pain of a burn can certainly not be classed beside a conviction, a surmise, a volition, etc., but beside sensory contents like rough or smooth, red or blue, etc.[9]

And Husserl continues at the same page:

> Every sensory feeling, e.g. the pain of burning oneself or of being burnt, is no doubt after a fashion referred to an object: it is referred, on the one hand, to the ego and its burnt bodily member, on the other hand, to the object which inflicts the burn. In all these respects there is conformity with other sensations . . . And though this reference is realized in intentional experiences, no one would think of calling the referred sensations intentional.

Thus far on the terminology regarding experiences in general and intentional experiences in particular. We shall look now into that part of Husserl's theory which is related to what we have referred to as Brentano's third elucidation of his definitional thesis, according to which there are various kinds of references to an object, and the distinctions between these kinds express important features of the structure of consciousness. Husserl adopts this idea and

elaborates upon it most resourcefully. The first impression is that he provides all the tools needed for an *extended theory of kinds of intentionality*. He writes:

> The act, e.g. corresponding to the name 'the knife on the table' is plainly complex . . . the knife is not the primary or full object of the judgement, but only the object of its subject. The full and entire object corresponding to the whole judgement is the *state of affairs* judged: the same state of affairs is presented in a mere presentation, wished in a wish, asked after in a question, doubted in a doubt, etc. . . . The state of affairs must obviously not be confused with the judging of it, nor with the presentation of this judgement. I plainly do not wish for a judgement, nor for any presentation . . .[10]

What is the most important aspect of these considerations for our analysis? The examples suggested by Husserl imply that we can speak of the self-same object as the object of more than one act. What I mean is not a multiplicity of acts identical in content, but such that occur at different times and/or in the consciousness of different people. What we are interested in are acts with different kinds of reference to the same object. *S*, in whose consciousness the presentation of the knife-on-the-table takes place, forms a judgement expressed by the indicative sentence 'the knife is on the table'. The same state of affairs obtains also as the object in *S*'s erotetic act expressed by the interrogative sentence 'Is the knife on the table?'. This applies also to *S*'s act expressed by the optative sentence 'I wish the knife were on the table'.

At this juncture Husserl introduced a distinction which proved to be of crucial importance in the subsequent development of the theory of intentionality. The distinction is between the 'content' or 'matter' of an intentional experience and its 'quality'. Husserl writes:

> We now turn . . . to another extremely important, seemingly plain distinction . . . This is the distinction between the general act-character, which stamps an act as merely presentative, judgemental, emotional, desiderative etc., and its 'content' which stamps it as presenting *this*, as judging *that* etc. etc. The two assertions '$2 \times 2 = 4$' and 'Ibsen is the principal founder of modern dramatic realism', are both, *qua* assertions, of one kind; each is qualified as an assertion, and their common feature is their *judgement-quality*. The one, however, judges one content and the other another content. To distinguish such 'contents' from other notions of 'content' we shall speak here of the *matter* (material) of judgements. We shall draw similar distinctions between *quality* and *matter* in the case of all acts.[11]

Although matter cannot be separated from quality within the intention, the quality may be identified by abstracting it from the matter, and *vice versa*. Furthermore, regarding their quality, all intentional experiences divide into a series of different *genera* or categories of intendedness toward an object. What is the nature and status of these categories? Is their number finite, can they be surveyed and explored? And if so, do they constitute a system with definite relations between its members? Is the status of each of them established once and for all, or is this great family of kinds of intentional quality perhaps in a state of perpetual flux? And above all, how are we to identify these categories?

I am asking these questions explicitly, in a *post factum* manner. There was no such calm attitude for Husserl who felt driven to go on with his theory, as if the tools for doing so were at hand. He writes:

> Content in the sense of 'matter' is a component of the concrete act-experience, which it may share with acts of quite different quality. It comes out most clearly if we set up a series of identical utterances, where the act-qualities change, while the matter remains identical. All this is not hard to provide. We recall familiar talk to the effect that *the same content* may now be the content of a mere presentation, now of a judgement, now of a question, now of a doubt, a wish etc. etc. . . . To be alike in 'content', while differing in act-quality has its visible grammatical expression; *the harmony of grammatical forms points the way to our analysis*.[12]

Having spelt out the crucial distinction between the 'matter' and the 'quality' of an intentional experience, Husserl advanced a striking working hypothesis: grammatical forms can be used to identify and explore the various kinds of intentional quality.

Was this hypothesis actually used in Husserl's inquiry? The answer to this question is a qualified 'no'. There were at least two preconditions for its effectiveness: (a) the grammatical

form must in fact be closely correlated with the intentional experience; and (b) philosophy of language should have been sufficiently developed at that time to provide a conceptual apparatus for exploring the various kinds of correlation between intentional experience and linguistic expression. The main reason for the hypothesis remaining unused was probably the non-fulfilment of the second precondition. But there was another, more specific reason for it, connected with the so called Aristotle-Bolzano controversy on the (logical) character of sentence-forms, which greatly occupied Husserl at that time.

The controversy may be rendered as follows: Granted that the mental act of judging is what confers meaning on an indicative sentence (when used to make a statement), we may ask whether for example an interrogative or optative sentence will also get its meaning directly from its correlated intentional experience. Aristotle's answer to this question would apparently have been in the affirmative. According to him there is no fundamental difference in this respect between judgement and other kinds of mental acts. Each of them confers meaning on the sentence form correlated with it.

At a certain stage Husserl was inclined to adopt this position, which would have coincided beautifully with his conception of the kinds of intentionality, which has been outlined above. Alas, towards the end of the *LI* he had changed his mind and adopted Bolzano's stand, as expressed in the latter's *Wissenschaftslehre* (Vol. I, par. 22): "A question like 'what is the relation of the diameter of a circle to its circumference?' asserts nothing about what it enquires into, but it asserts something nonetheless: our desire, in fact, to be informed concerning the object asked about. It is indeed capable of truth and falsehood, and is false when our desire is misstated by it".[13]

Although Husserl objects to the *way* in which Bolzano reduces the non-assertives to assertives, he adopts Bolzano's position in principle, suggesting his own idea on the relation between the kinds of acts and sentence-forms. An illuminating passage in this respect goes like this:

> Talk of 'expressing' is . . . ambiguous, and it remains so even when we connect it with the *acts* to be expressed. What are expressed may be, on the one hand, said to be sense-*giving* acts, to which in the narrower sense, 'voice' is given. But there are other acts which can also be said to be expressed, though . . . in a different sense. I refer here to the very frequent cases in which *we name acts we are now expressing*, and through such naming manage *to say that we are experiencing them.* In this sense I 'express' a wish through the words 'I wish that . . .', a question through the words 'I am asking whether . . .', . . . and so on. Naturally, we can pass judgement on our own inner experiences, just as we can pass judgements on outward things, and when we do the former, the meanings of the relevant sentences will reside in our *judgements* upon such experiences, and not in the experiences themselves, our wishes, questions, etc.[14]

Thus Husserl found a way (or so he thought) preferable to Bolzano's method to reduce the non-assertive to assertive forms, and this accomplished, the issue was settled for him for the time being, as is explicitly declared at the end of the VI[th] Logical Investigation.[15]

It would not be correct to say that as a result Husserl lost interest in the non-assertive forms and their correlates on the intentional level. He mentioned them on several occasions in his later writings. In *Erfahrung und Urteil* we even find quite an extensive discussion on the interrogative.[16] But, as far as I know, he never attempted a systematic treatment of these forms.

III

As regards this issue, I would like to substantiate the view that a continuity is detectable from Husserl's to Searle's position, even if Searle may not have been aware of it. As a matter of fact, he ascribes to himself a deviation from what he calls 'the tradition' of the theory of intentionality, and sets out to tell us in what ways his use of the term 'intentionality' differs from the 'traditional' one.[17] We should not, however, take this last claim too seriously. In general, when it is correct, the deviation is not very significant, and when the deviation seems of importance, the allegation is not quite accurate. For example, Searle argues that his use of

the term 'intentionality' is non-conformist, as it stresses the difference in scope between intentionality and mentality: there are mental states which are not intentional, though not *vice versa*. This is a significant contention, but as we have seen, it was Husserl's position as well.

Further, Searle rejects calling the intentional state 'an act'. There are indeed, according to him, mental acts, as for instance "doing arithmetic in your head or forming mental images of the Golden Gate Bridge", which are also intentional states, but there are also countless intentional states which are not acts, like desiring, hoping, fearing, believing. An act is something that you perform, and in those other states none of it is occurring.[18]

We can reply to this that Husserl did indeed use the term 'act' as synonymous with 'intentional experience' (which Searle labels 'Intentional state' or 'event'), but he had explicitly made it a *technical* term. At a later stage of the development of his doctrine, when he came to conceive of the intentional experience as constitutive of its objects, the concept of 'act' came close to the sense of acting or bringing about, but even then Husserl preserved the distinction between this aspect of intentionality and acting as purposeful intervening in the course of events. The term 'act' seems to have remained for Husserl a technical term all along the way.

Searle's third contention about deviating from tradition concerns the relation between the term 'intention' in its ordinary, pre-analytic sense, and 'intentionality' as that property of mental states or events which is the subject-matter of our inquiry. Searle argues for a clear separation of the two. The term 'intention' in its pre-analytic sense is essentially connected with acting, speech being also considered a kind of acting. We form an intention to do something in the far or near future, or we do something with an intention to reach a certain goal, or to make something happen. Such an action will be called 'intentional', but there is no point, says Searle, to consider a belief, or a hope, or a desire 'intentional' in this sense. The mental states of belief, hope, or desire, imply a directedness toward some object or some state of affairs, and are thus identified as intentional, though no action is involved. Hence, we should (a) avoid using the noun 'intention' for those intentional states which have nothing to do directly with action; and (b) to specify the expressions 'intentional' and 'intentionality' as technical terms. Accordingly, every intention-for-doing-something and every intention-in-action is an intentional state, but not *vice versa*.

Here we can certainly see a deviation from the terminology used by Husserl and his followers, but hardly anything beyond terminology is involved. Indeed, Husserl often used the term 'intention' as synonymous with 'intentional experience' or with 'act', but similarly to Searle he took care to distinguish between the 'normal' or 'every day' use of the term, as essentially connected with acting, and the technical use of it signifying the subject-matter in the theory of intentionality. In German there is no great danger of misunderstanding, as there are different words for the ordinary sense of intention (*Absicht, Vorsatz*). Moreover, this technical use of 'intention' points to the etymological connection with the mediaeval tradition, from which it was taken by Brentano, as mentioned before.

Let us proceed now to Searle's theory itself. It could be said that the working hypothesis which Husserl had suggested, but failed on the whole to implement, is now applied by a masterly hand. The 'harmony of grammatical forms' as a guide in the analysis of intentionality received a radical interpretation in Searle's procedure of exploring the structure and logical properties of the intentional state by means of the model of a speech act.

This indeed is the *main* procedure that Searle has adopted to elaborate his theory. It is, however, underlied by the following, somewhat 'ideological' assumption, the importance of which cannot be overestimated.

The procedure became possible, since the model could be based on solid knowledge in the field of Philosophy of Language in general and on Searle's theory of speech acts in particular. Nevertheless, one must not jump to the conclusion that what Searle does is interpreting intentionality in linguistic terms. It is only an empirical fact that historically the development of the Philosophy of Language proceeded at a faster pace than did the inquiry into the nature of intentional states. As fas as the logical relation between these fields or layers is concerned, the intentional states come first. As Searle puts it: "Language is derived from intentionality

and not conversely. The direction of pedagogy is to explain Intentionality in terms of language; the direction of logical analysis is to explain language in terms of Intentionality".[19]

In the course of his analysis Searle indicated a number of parallelisms between the intentional state and the speech act, and doing so he had developed a very useful conceptual apparatus for a systematic treatment of problems of the theory of intentionaltity and related topics.

I must rest content with analysing only a few of these parallelisms and connections, the ones which are particularly relevant to our discussion. First and foremost let us consider the parallelism between two distinctions, each of which is basic in its field: one of them is between the 'propositional content' and the 'illocutionary force' of the speech act, and the other is between the 'intentional (or representational) content' and the 'psychological mode' of the intentional state. What is immediately noticeable is that on the side of the intentional state the distinction is the same as the one we have found in Husserl's conception and described as one of the most important single findings in the field. Searle writes in this connection:

> Just as I can *order* you to leave the room, *predict* that you will leave the room, and *suggest* that you will leave the room, so I can *believe* that you will leave the room, *fear* that you will leave the room, *want* you to leave the room, and *hope* that you will leave the room. In the first class of cases, the speech act cases, there is an obvious distinction between the propositional content *that you will leave the room* and the illucutionary force with which that propositional content is presented in the speech act. But equally in the second class of cases, the Intentional states, there is a distinction between the representative content *that you will leave the room*, and the psychological mode, whether belief or fear or hope or whatever, in which one has that representative content.[20]

The relationship between the two levels is further elucidated by another parallelism between them, this time with respect to the so called direction-of-fit. As much as there are speech acts with the direction of fit of word-to-world, as for example a *statement* or a *description*, and speech acts with the direction of fit of world-to-word, like an *order* or a *promise*, so there are also intentional states with the direction of fit of mind-to-world, as for example *belief* or *hope*, and intentional states with the direction of fit of world-to-mind, like *desire* and *intention* (in its pre-analytic sense).

Finally, we are given a powerful key-concept of the 'conditions of satisfaction', which functions parallelly at both levels, linguistic and mental alike.

> We say, for example, that a statement is true or false, that an order is obeyed or disobeyed, that a promise is kept or broken. In each of these we ascribe success or failure of the illocutionary act to match reality in the particular direction of fit provided by the illocutionary point. To have an expression, we might label all of these conditions 'conditions of satisfaction' . . . [T]his notion of satisfaction clearly applies to Intentional states as well. My belief will be satisfied if and only if things are as I believe them to be, my desires will be satisfied if and only if they are fulfilled, my intentions will be satisfied if and only if they are carried out.[21]

The concept of conditions-of-satisfaction proved to be extremely fruitful at later stages of the analysis, when Searle used his conceptual apparatus to inquire into phenomena like sense-perception, activity, the relation between intensionality and intentionality, proper names, etc.

There is a considerable overlap between this concept of conditions-of-satisfaction and Husserl's distinction between intention and fulfilment. As we read in the *LI*:

> We may now further characterize the consciousness of fulfilment by seeing in it an experiential form which plays a part in many other fields of mental life. We have to think of the opposition between wishful intention and wish-fulfilment, between voluntary intention and execution, of the fulfilment of hopes and fears, the resolution of doubts, the confirmation of surmises, etc., to be clear that essentially the same opposition is to be found in very different classes of intentional experiences.[22]

These remarks on what Husserl calls the 'wider class of experiences' appear again in the context of the Sixth Investigation, as a kind of digression from the main subject, which is the phenomenological elucidation of knowledge'. The distinction between intention and fulfil-

ment is investigated in close connection with only one class of acts, the so called 'objectifying acts' of judging and asserting.

I have thus far outlined the common areas of the two approaches without specifying how far does the overlap extend. This, I believe, is a problem deserving a separate treatment.

IV

To sum up: the context of our analysis in its historical perspective is a restricted one. The point of its departure is Brentano, but it is only the Brentano of the *Psychologie vom empirischen Standpunkt*, where he proposed his initial thesis on intentionality as the *differentia specifica* of mentality. At the core of this thesis we have found two distinctions intersecting each other: one between the intra- and the extra-mental existence, the other between intendedness and the lack of it. We have also understood that at least two problems cropped up from that terse formulation to challenge all who would proceed along Brentano's path: the problem of the ontological status both of the object (or the state-of-affairs) contained in the mental phenomenon and of the mental phenomenon containing it; and the problem of the variety of kinds of intentionality.

Husserl, and here we mean Husserl of the *LI*, took up the challenge of both problems; however, with respect to our discussion, the second one expresses better the continuity between the respective positions of the two thinkers. In pursuing this line of thought Husserl produced an important analytical tool: the distinction between the Matter (or Content) of the intentional experience and its Quality. He also proposed a device to guide us in identifying and analysing those *kinds of intentionality*: use of the alleged correlations between linguistic forms and intentional experiences. Unfortunately, this ingenious device had to remain, at least at the stage under consideration, unused. Even at later stages of the history of Husserl's theory of intentionality this particular part of it has remained underdeveloped, although there is some textual evidence that he always meant to come back to it.

At this point we have found it possible to speak of continuity between Husserl's position and that of Searle's. We have called Searle's procedure of using the speech-act as the model for analysing the logical properties of the Intentional state 'a radical interpretation of Husserl's working hypothesis' and have pointed out that Husserl's Matter and Quality, in a slightly changed terminology, gained the rank of key-concepts in Searle's theory. In fact, we have suggested, the parallelisms between the two positions go much further.

Still, this applies only to a comparison within the restricted context, as described above. Should this context be broadened, far reaching differences will emerge. For example, Searle says explicitly in his *Intentionality* (p. 160) that he has adopted a "resolutely naturalistic" approach, that he thinks of Intentional states, processes and events, as a part of our biological life history, in the way "that digestion, growth and secretion of bile are part of our biological life history". The later Husserl, Husserl of the 'phenomenological reduction' and 'intentional constitution' would have vigorously rejected such an attitude, as contradicting his 'transcendental idealism'.

Here, it seems, we are facing a real choice. To consider it, however, we would have to dive into deep philosophical waters, and such an undertaking must await another opportunity.

ENDNOTES

[1] Brentano, Franz, *Psychology from an Empirical Standpoint*, ed. Linda L. McAlister (London 1973), p. 88f.
[2] In McAlister's edition: 'a thing', in the original: 'eine Realität'.
[3] Rarely mentioned by commentators.
[4] Cf. J. N. Mohanty, *The Concept of Intentionality* (St. Louis, Missouri 1972), on the quoted passage in general, but particularly on this point, p. 6.
[5] McAlister's edition, p. 88n.
[6] On the history of this concept see W. & M. Kneale, *The Development of Logic* (Oxford 1962), p. 224ff.; I.M. Bocheński, *A History of Formal Logic* (Notre Dame 1961), p. 154ff.
[7] See e.g. the Marras-Spiegelberg controversy, as in Linda L. McAlister (ed.), *The Philosophy of Brentano* (London 1976), pp. 108–139.
[8] Quoted from Findlay's translation, as in Edmund Husserl, *Logical Investigations*, (London 1970), Vol. II (hereafter *LI*), p. 569f.
[9] *ibid.*, p. 572f.
[10] *ibid.*, p. 579.
[11] *ibid.*, p. 586.
[12] *ibid.*, p. 586, italics added.
[13] Quoted from *LI*, Vol. II, p. 840.
[14] *ibid*, p. 677. For a critical analysis of these considerations of Husserl's see J. N. Mohanty, *Edmund Husserl's Theory of Meaning* (The Hague 1964), pp. 80–86, as well as his "Husserl's Theory of Meaning" in F. Ellison and P. McCornick (eds.), *Husserl's Expositions and Appraisals* (Notre Dame 1977), pp. 26 and 34.
[15] *LI*, Vol II, p. 849ff.
[16] Husserl, Edmund, *Erfahrung und Urteil: Untersuchung zur Genealogie der Logik*, ed. Ludwig Landgrebe (Hamburg 1948), pp. 371–380. For an extensive survey and analysis of the various phases in the development of Husserl's conception of intentionality see Mohanty, *op. cit.*, Part II.
[17] See John R. Searle, *Intentionality: An Essay in the Philosophy of Mind* (Cambridge 1983), p. 1. See also "Von der Sprechakttheorie zur Intentionalität: ein Gespräch mit John R. Searle" in *Information Philosophie* (Basel 1984), p. 24ff.
[18] Searle, John R., *Intentionality*, p. 3.
[19] *ibid.*, p. 5.
[20] *ibid.*, p. 5f. Jitendranath Mohanty was probably the first who noticed the striking coincidence of these distinctions. See his "Husserl's Theory of Meaning" mentioned in Footnote 14, above, p. 22n. (quoted by Hubert L. Dreyfus in the text referred to below, Footnote 22).
[21] Searle, J. R., *Intentionality*, p. 10.
[22] *LI*, Vol. II, p. 699. Prof. Dreyfus arrives at a similar assessment of the relation between these basic distinctions in the Introduction to his collection *Husserl, Intentionality and Cognitive Science* (Cambridge, Mass., 1982), p. 4ff. Dreyfus' text came to my attention after the present paper had been submitted.

* * *

RE-EMERGENCE OF THE CONCEPT OF CONSCIOUSNESS IN TWENTIETH CENTURY SCIENCE AND PHILOSOPHY

Rudy Krejci
University of Alaska, Fairbanks

INTRODUCTION

> Never in the history of mankind were there so many people shaped and influenced for such a long period of time by rather a few profound ideas as during the last four centuries.[1]
>
> Anonymous

During the last four hundred years and into our century, a series of conceptual revolutions have taken place forcing us to change our way of thinking. The Copernican cosmological revolution destroyed the millenia old belief in the central position of the earth in the universe which was so important for the concept of a God-created universe and of man in it. A new emerging view of the nature of the universe produced a synthesis of physics and astronomy culminating in the classical physics of Newton which gave us a picture of the world that remained the same no matter how and from where one looked at it. It was a definite picture, a mechanistic view reflecting a mechanistic reality. This change in the concept of the universe implied a change in the concept of man. During the 18th century a new science of man and of society was formulated, patterned after physics and producing social physics, social mechanics, social statistics, social energetics, associanistic psychology, etc. During the 19th century four other revolutions followed within the mechanistic paradigm: first, the Darwinian theory of evolution, destroying the concept of a universal design with man on the top, and introducing the concept of survival of the fittest as a new mechanistic selector in the development of species. This was followed by Karl Marx explaining the history of human societies as a later dialectical development of cosmic matter. Materialism was regarded as the safest foundation of science and, correspondingly, social science could be nothing but some kind of economical determinism. Similarly, Freudian psychoanalysis projects the existing mechanistic paradigm upon the study of personality, declaring man as not being in charge of himself but rather a victim of unconscious drives and two forces: the self-preservative and the sexual drives whose primary interests are the optimal satisfaction of both of them. Thus man, unable to control himself, became a true object of scientific studies together with other species as a stimulus-response aparatus as seen in the emergence of Pavlovian teaching which prepared the way for the triumph of Behaviorism in the 20th century.

Newtonian science, due to its prestige, provided a new scientific paradigm, which in due time, produced a corresponding social science and science of man. In this model, science was concerned only with quantities and measurements for these were the primary qualities of the objective world which were captured in linguistic replicas and mathematical formulations. Ethical and moral issues were not scientific and therefore of no concern. One talked about the world without the awareness that someone was doing the talking. The phenomena of mental events, minds, psyche, self, disappeared in the process of studying man scientifically, consciousness became in these scientific circles "a buzz-word of our flower generation" and as such, of no scientific concern. The concept of man that emerged held that the human mind was a passive, blank sheet of paper, which in turn became the foundation of a political system of democracy based on the equality of social atoms at the time of birth. Thus the mechanistic concept of the universe produced a predominantly mechanistic society, with mechanistic man, reared in a mechanistic culture, medicine, law, education, economics, industry etc.

While all of the previous have been and to great extent still are going on, a new tacit conceptual revolution was in the making, finding at first its expression in the foundations of mathematics, thus producing new geometries with new implications for the whole realm of human thinking. Hand in hand with the radical changes in the foundations of mathematics went new modifications of the millenia old Aristotelian system of logic. The changes here were small but they were not expected to happen in a field regarded to have been definitely stated such a long time ago. By the end of the 19th century there were new ways expressed in discontinuities in literature, atonality was introduced into music, cubism represented similar innovation in art. Political, social and economical crisis accompanied this time of transition.

When at the turn of the century, classical physics found itself in a serious crisis, two new theories of modern physics (Einstein's theory of relativity and quantum mechanics) introduced a new conceptual reconstruction of reality, revealing the world we cannot perceive with our senses or adequately describe in our language. Time and space are conceived as inseparably connected, matter turns into fields of energy losing its substance and the consciousness of the observer-participator becomes an integral catalyst of the pattern of interactions of the whole. Where one would expect a fundamental agreement among the leading exponents of the new way of thinking one must be disappointed facing a veritable schism between Einstein on one hand and Bohr, Heisenberg and other quantum physicists on the other hand. What is the meaning of the conflict, and is the conflict basically of a philosophical, scientific or religious nature, or are all of them just three aspects of the same? Who is here to say, who is here to judge?[1]

The twentieth century will certainly occupy a very distinct role in the history of humanity due mostly to its multifaceted contradictions shaking the foundations of entire societies, furthermore by its atrocities, mass murders, ideological fanaticism hand in hand with the new emerging concept in the physical sciences and their technologies, bringing mankind to the brink of extinction. At the same time, bringing the people of the earth into close physical proximity and into better understanding. We can justly say at the close of our century that we have lived in a century of revolutions which for the first time involved the entire earth with its populations indicating the necessity of a new cultural, political and socal synthesis.[2]

I

> The only goal of science appeared to be analytical, i.e., the splitting up of reality into ever smaller units and the isolation of individual causal trains.[3a]
>
> Ludwig von Bertalanffy

The purpose of this essay is to explain the re-emergence of the concept of consciousness in 20th century science and philosophy. First, since the re-emergence is claimed to have occurred, I have to deal with the very existence of the concept of consciousness, how it emerged originally, how it was treated in the past and how it submerged later on, before it re-surfaced again in our own times.

Consciousness-self-knowledge, self-awareness, intentionality, introspection—is basically an art or act of observing one's own mental states and processes with the purpose of achieving self-knowledge—a phenomenon known to all people of all cultures and times, a human universal creating the same individual way of successfully achieving it. In our Western tradition Socrates is the symbol of a true seeker of wisdom (know thyself), he is the light in the darkness, and as Plato describes it in his Republic, there is a way of achieving the Good, the Absolute in terms of which the mind can recognize its own ideas. After training the body and the mind for many years a conscious state of harmony can be achieved, which in turn can introduce the student into the study of dialectics, where ideas are studied in their relationship away from all sensory imports till one reaches the level of the mind which is considering only the unchangeable Absolute Good itself.[3] Plato's philosophy survived in the Academy and in Neo-

platonism, finding new formulations in Plotinus and later in St. Augustine who provided the first masterful synthesis of the ancient quest for wisdom within Christian teaching. Man's search for absolute certainty was now defined in religious terms and so it stayed till the end of the Middle Ages.

Modern philosophy begins in the West with Descartes who continued the search for absolute knowledge and found it in his famous Cogito-consciousness. He assumed that philosophy should start with the study of consciousness, however, this very claim did not find an expected follow up. His conscious mind was for him nothing but a mirror reflecting the world. He even admitted that the mirror might distort the world, however, he was convinced that man could never know for sure if this was the case because his mind can never get "out there" and check on it.[4] (4) Descartes never imagined that the mind itself could be an important variable in our conception of the world. The followers of Descartes philosophized out from his inconsistencies everything there was—so we had philosophies of dualism, monism, pantheism, idealism, and materialism claiming him as their legitimate father.

Among the British empiricists it was especially John Locke who maintained that the Self was a continuity of consciousness, independent of material bodies and of substances as well. Locke's great follower, David Hume, became the most influential philosopher in Anglo-Saxon thinking up to our present time. Hume, trying to find the mysterious self of Descartes, was not successful, what he found was only a "bundle of impressions", no general, overriding self, no unifying self of awareness. Did Hume not notice that someone was doing the looking? Evidently it never occurred to him. After his failure he hoped that others might be luckier in the future in solving the problem.[5] The followers of Hume, like William James, who agreed with him, maintained that introspection does not reveal an overriding self. During the 20th century, the Anglo-Saxon tradition put Hume's argument into modern logical gown in Gilbert Ryle's "systematic elusiveness of the concept of 'I'".[6] The American behavioristic school of Watson and Skinner boldly proclaimed that there were no mental events, not being aware, of course, that it followed.

Kant agreed in principle with Hume, that while a source to thinking must be thought, one can never know it. The self, the "I" is beyond an experience, unknowable.[7] Modern existentialism found in Sartre even a more radical position stating that cansciousness itself must be nothing, non-being: Therefore, there is no true self, no unifying sub stratum, no real source for our thoughts.[8]

In the view of most modern Western thinkers, the original Greek quest for self-knowledge, wisdom, is not attainable. The Delphic maxim—Know Thyself—is an illusion which cannot be reached. It appears that Plato's dialectic technique was lost, Descartes did not give us the source of conscious thought, introspection reveals only individual bundles of impressions, and consequently, consciousness turns necessarily into an empty concept—an unnecessary philosophical aether . . .

II

> Neurophysiologists will not likely find what they are looking for outside their own consciousness, for that which they are looking for is that which is looking.[9a]
>
> Keith Floyd

The re-emergence of the concept of consciousness during 20th century development can be traced to three major sources: 1. philosophical, 2. scientific, 3. Oriental philosophical imports in the West, and combinations of these three. The first source is represented by the school of phenomenology and of its founder Edmund Husserl (1859–1938), whose ideas have exerted a powerful influence on European and to a certain extent also on American thinking between the two World Wars. The phenomenological method has been widely applied outside of the field of philosophy where it has made new contributions in the fields of literature, art, aesthetics, social sciences, counseling, pedagogy, religion and many others. The central concept of

phenomenology and the soil it germinated came from Franz Brentano's notion of intentionality. The mind, or conscious mind, or consciousness is always intentional, never merely reflecting the world but producing its own distortions due to our various mental positions—emotions, idiosyncracies, likes, dislikes, prejudices, etc. Consciousness itself cannot be taken for granted, furthermore, it has to be investigated. Husserl defines phenomenology as the study of the structure of consciousness, which by being necessarily intentional is at the same time prejudiced. Scientific pursuits can partially help to eliminate some of them, however, there are also deeply seated habits to our thinking and perception. Science itself may be a mere product of them. Philosophy, to become a true exact science, has to fully recognize that there is a large field inside of man himself which has to be fully studied and comprehended before anything else.[9] Up to now, consciousness has become more or less a hostage of external objects, and thus enslaved it had to become passive, servile, and gutless.

Phenomenology is predominantly a method and it must follow that each application will happen to become necessarily highly individualistic as the followers of Husserl have shown so well. Phenomenology is perhaps, another name for science, a true science, the results of which cannot be petrified. Phenomenology with its "going back to the things themselves" is the direct study of consciousness in order to gain an increased control of our own existence. As a cognitive procedure, the phenomenological method consists fundamentally of an intellectual observation of the object based on intuition. In order to achieve that, a threefold reduction, or exclusion, or bracketing out, (epoché)—is required: 1. one has to bracket out one's own beliefs in the real existence of the object; all subjectivity has to be set aside, 2. all theoretical knowledge, hypotheses, theories, and proofs derived from other sources, have to be set aside. Only the given will be admitted. 3. all the traditions, and what others taught about the object before have to be excluded. The object under investigation has to be subjected to a further two-fold reduction. Its existence has to be disregarded and everything inessential has to be omitted. Thus, the successful steps of bracketing reveal the naked phenomenon in front of the investigator. And here is exactly where the method ends—the conditions are satisfied for the creative part which will be carried on intuitively. Because man is not a syllogism—nothing has to happen! to-day, maybe to-morrow—it cannot be commanded when the creative charge is to happen, or if it will happen at all. The negative aspect of the phenomenological method represents a deep insight into human creativity.

Husserl came to think that the intentional consciousness represented an intersubjective world,

a true community of persons living together in the prenatural, prescientific universe—the task of phenomenology will be the exploration of the human-life-world. (Lebenswelt)[10]

Phenomenology is a veritable Copernican revolution, fundamentally changing the entire realm of human thinking. Awareness of the continuous intentionality of consciousness is then responsible for a continuous process of change. In this state, phenomenology certainly establishes a certain control of continuous awareness which seemed to be totally lacking when mind-changing drugs were used during the last twenty years of experimentation: in these states the penetrating insights may be more intensive, but, without a context of lasting awareness.

All in all, phenomenology has inspired new developments in the understanding of consciousness and with its aesthetic, ethical and mystical experiences supplies a new starting point for mankind at the end of the 20th century.

III

> The common division of the world into subject and object, inner world and outer world, body and soul is no longer adequate and leads us into difficulties.[11a]
>
> Werner Heisenberg

Something rather similar to phenomenology can be observed in the radical shift in perspec-

tive which took place in modern physics when Einstein, Planck, Bohr and Heisenberg and their many followers realized that the completely passive observer never existed in scientific investigations. What were the implications in the past when Galileo talked about the primary and secondary qualities, and the great Newton did not pretend to give us a hypothesis, and Comte once and forever pontificated in his "laws of three stages" the triumph of positive science over metaphysics and religion? Science was always very progressive in fighting any kind of dogmatism save its own. It is apparent now that the classical world-view was identical with the belief that the universe was a huge aggregate of dispersed homogeneous matter changing its configurations according to the immutable laws of mechanics resting on classical concepts of space, time, matter, motion, causality, determinism and our pictorial character of knowledge. The very transition from the old to the new way of thinking is best characterized by words—picture with its passivity, as used in the past and conception with its creativity, as we use it presently.[11] The 20th century is undergoing perhaps the most engaging and deepest epistemological metamorphosis in the entire history of humanity. The key term that the new physics will talk about is the discovery of the observer and what it is that he really observes: To this, relativity theory answered—relations, quantum mechanics—probabilities.[12] The observer is, of course, a conscious observer, who cannot observe without altering what he observes—both of them are involved in an inseparable interrelation. We can very well conclude that what we experience is not an external objective reality as we have believed in the past, but our direct interaction with it. As Heisenberg said: "What we observe is not nature itself, but nature exposed to our method of questioning."[13] Our answers can only be understood in the context of their inseparable questions. Furthermore, our language does not picture our more advanced thoughts when we deal with the objects not directly observable. We do not even have to assume an objective reality apart from our experience. We deal with the microcosmos and macrocosmos in continuous process which interact, in our consciousness. In this dynamic universe we can only describe the statistical behaviour of the systems and make predictions—our observations are constant fusions and alterations of the observed. The conceptual framework of the new physics with an emphasis on the observer and consciousness is responsible for creating a mental atmosphere that reminds one of deep mystical vision. The men responsible for the new physics, without exception, exemplified the ongoing metamorphosis of their own unprecendented historical search with their belief in a mystical world-view that embodies the world as a spiritual, rather than material phenomenon.

There was a remarkable pilgrimage to China made by Niels Bohr in 1932. "Bohr came back from his visit to China in 1932 inspired by the tradition of learning in China and filled with admiration for the great Chinese thinkers of past and present. The central point of quantum theory, Bohr already had been forced to conclude five years earlier, is complementarity . . . In the East, Bohr discovered to his happiness, that complementarity could be the natural way of thinking. To symbolize complementarity, Bohr chose the Chinese yin-yang symbol and composed the Latin motto—not contraria sed complementa—The way of thinking symbolized by complementarity is so important in his view that every school child should be taught it."[14] Almost 50 years later Prof. Wheeler himself was invited to come to China and present three lectures on history and on contemporary views of quantum theory. In his first lecture he stated to his Chinese audience: "We have learned in our day that it is wrong to think of the universe as—sitting out there—. In the older physics, it was natural to think of the observer as, in effect, looking at "existence" from behind the safe protection of a 20 cm slab of plate glass. In contrast, quantum mechanics teaches the direct opposite. It is impossible to observe even so miniscule an object as an electron without, in effect, smashing that slab and reaching in with the appropriate measuring equipment. Moreover, the installation of apparatus to measure the position coordinate, x, of the electron, automatically prevents the insertion in the same region at the same time, of equipment to measure its momentum, P; and conversely. Furthermore, the choice of what to measure—or whether to measure at all—has irretrievable consequences for what one has a right to say about the electron in the future."[15]

Wheeler thus promotes the observer to a direct participator.

Eugene Wigner, one of the living men responsible for the new physics stated: ". . . it is not possible to formulate the laws of quantum mechanics without reference to consciousness." He believes that there are two kinds of reality. "These are so different that they should have different names. The reality of my perceptions, sensations and consciousness is immediate and absolute. The reality of everything else consists in the usefulness of thinking in terms of it. This reality is relative and changes from object to object, from concept to concept."[16] We come here in contact with Wigner's strong belief in the existence of consciousness and the necessity of its scientific study.

It is remarkable that every major physicist involved in the work of the new physics was completely in favor of the interdisciplinary dialogue. Probably their deep involvement in this unique search into the mysteries of nature led them without exception into the mystery of human consciousness. Einstein, Bohr, Heisenberg, Schroedinger, De Broglie, Jeans, Planck, Pauli, Eddington and their followers have created for mankind what one day will become known as a New New Testament, a respectable sequel to the other two Testaments and all other great documents of both eastern and western humanity.[17]

IV

> . . . because our science—Greek science—is based on objectification, whereby it has cut itself off from an adequate understanding of the subject of Cognizance, of the mind. . . . our present way of thinking does need to be amended, perhaps by a bit of blood-transfusion from Eastern thought.[18a]
>
> Erwin Schroedinger

Our present day culture is in the midst of profound changes. After 400 years of energetically practising the study of the outward world in the West we have been able to reach an understanding of the physical and to some extent biological phenomena. On the other hand the East by undertaking the study of the inward world for thousands of years arrived at a deep understanding of the inner universe of consciousness. In our western modern science we do not claim to possess anything more than embryonic ideas about consciousness and its role. The traditional East, particularly in its old Vedic scriptures, together with the six derivative systems of Indian philosophy possesses a well developed understanding of consciousness. Mostly due to our modern technological innovations the former cultural separations were cancelled resulting in the present day mutual interchange of ideas and goods. There are many recent developments that have begun to extend the scope of the Oriental imports in the West. The audience, a recipient of Oriental influences, represents a rather motley group as a cross-section of our western society. There exists a "counter-cultural community," opposed to science, and showing a great distaste for all of the rational thinking and its products—western logic, machines, gadgets and computers and their applied technologies. Various Oriental philosophies and religions are enthusiatically embraced satisfying the demand for a nonlinear, arational, intuitive and highly personal way of thinking and living. Then we have groups with a widespread interest in mind-altering drugs, such as psychedelic drugs, creating addictions together with the promise of a panacea for many of our young disillusioned generation.

Among today's very popular teachings we have more developed traditions including Zen-Buddhism, Tibetan Buddhism, Sufism, various aspects of Yoga, and Transcendental Meditation which, in itself, represents the theoretical foundation of the science of creative intelligence. During the last few decades there has been a great revivial of interest in the study of consciousness within Western philosophy which is challenging the prevailing empirical position on the topic. Views of such philosophers like Butler, Reid, and McTaggart have been reintroduced, joined by new works of Chisholm, Madell, Ornstein and others.[18]

Another interesting phenomenon during the last 15 years has been the emergence of many studies written by scientifically trained writers who have attempted to link western science, especially physics, to oriental mystical traditions against violent objection by some phy-

sicists.[19] It is becoming clear that the notion of consciousness which has been totally absent from our western science, creates havoc in the conceptual structure of many contemporaries who have been trained in the old paradigms. A great part of our educational system in the West is still deeply oriented in the tradition and perpetuates the stubborness of the old regime. In spite of much progress in Western education we can still regard it more or less Ptolemaically oriented—under culture we understand western culture, under science western science, under art western art, under logic western logic, under thinking western way of thinking, under values western values. It will require a lot of time before we are freed from our narrow-minded provincialism and prejudices. In spite of all of the contemporary oriental imports in the West, most of them have achieved a rather superficial fashionable effect, many of them are turned into our typical western band-waggons. They come and go. There was an insightful work written during the second World War by F. S. C. Northrop—*Meeting of East and West* which was intended by the author to be a preparatory work introducing comparative cultural values of the post-war world;—looking back almost forty years, we can see that the work has hardly made a dent in our cultural consciousness. In spite of all the defficiencies our present human nature is responsible for, both the value systems of the East and the West must be viewed as being complementary and not contradictory as it has been understood by both sides up to now. The Western revival of interest in the spiritual and the Eastern interest to improve their material well-being is such a complementary enterprise requiring mutual help based on respect and understanding.

ENDNOTES

1a Anonymous

1 Capek, Milic, *The Philosophical Impact of Contemporary* Physics (Princeton, N. J., 1961), pp. 3–6, 361–399.

2 Leshan, Lawrence and Margenau, Henry, *Einstein's Space and Van Gogh's Sky—Physical Reality and Beyond* (New York 1982), pp. 3–41.

3a Bertalanffy, Ludwig von, *General System Theory* (New York 1968).

3 Plato, *Republic* (Oxford Univ. Press, 1971), Chapter IX, 376E–392C, ibid, VI. 509b–518d.

4 Descartes, René, *Discourse on Method and Meditations* (The Library of Liberal Arts 1960), Second Meditation, pp. 81–91.

5 Hume, David, *Treatise of Human Nature* (London, J. M. Dent, 1951), Vol. 1, pp. 238–244.

6 Ryle, Gilbert, *The Concept of Mind* (University Paperbacks, Barnes and Noble, New York), pp. 245–266.

7 Kant, Immanuel, *Critique of Pure Reason* (Norman Kemp Smith's), pp. 83–89, 132–169.

8 Sartre, J. P., *The Transcendence of the Ego* (Noonday press, New York), pp. 80–102.

9a Floyd, Keith, *Of Time and the Mind*, "Fields within Fields", No. 10 (1973–74).

9 Husserl, Edmund, *Ideas: A General Introduction to Pure Phenomenology* (London and New York 1931), per 75.

10 Husserl, Edmund, *Die Krisis der europäischen Wissenschaften und die transzendentale Phänomenologie* (the Hague 1954), pp. 347–on.

11a Heisenberg, Werner, *Physics and Philosophy* (New York 1958).

11 Capek, Milic, pp. 3–4.

12 Eddington, Arthur, *The Philosophy of Physical Science* (Ann Arbor Paperbacks, 1958), Chapter V, pp. 70–105.

13 Heisenberg, Werner, (1958), pp. 58–on.

14 Wheeler, John Archibald, *Physics and Austerity—Law Without Law—Workings Paper* (University of Texas, 1982), pp. 1–2.

15 Wheeler, John Archibald, *Physics and Austerity—first lecture—This Incredible Quantum Business*, p. 9.

16 Wigner, Eugene, *Symmetries and Reflections* (Bloomington, Ind., 1967), pp. 210–288.

17 Wilbur, Ken, *Quantum Questions—Mystical writings of the world's great physicists* (New Science Library 1984), pp. 3–208.

18a Schroedinger, Erwin, *What is Life and Mind and Matter* (Cambridge Univ. Press).

[18] Kupperman, Joel J., "Investigations of the Self", *Philosophy East West* (Jan. 1984), pp. 37–49.
Nalimov, V. V., Three important contemporary translations from Russian: *Faces of Science* (Philadelphia, Pa, ISI Press, 1981); *In the labyrinths of language: a mathematician's journey Philadelphia,* P, ISI Press, 1981; *Realms of the Unconscious: The Unconscious: The Enchanted Frontier* (Philadelphia, Pa, ISI Press 1982), pp. 1–19; 140–146; 257–300.

[19] Capra, Fritjof, *The Tao of Physics* (Shambala 1975).
Capra, Fritjof, "Modern Physics and Eastern Mysticism", *The Journal of Transpersonal Psychology* 8 (1976), pp. 20–39.
Pagels, Heiz, The Cosmic Code (New York 1982).
Bohm, David, *Wholeness and the implicate order* (London 1980), pp. 196–213.
Heisenberg, Werner, *Physics and Philosophy* (New York 1958).
Heisenberg, Werner, *Physics and Beyond* (New York 1971).
Heisenberg, Werner, *On Modern Physics* (New York 1961).
Sarfatti, J., "The Physical Roots of Consciousness" in Mishlove, *The Roots of Consciousness*, New York 1975, pp. 279ff.
Talbot, Michael, *Mysticism and the new Physics* (A Bantam Book, 1981), pp. 1–64.
Prigogine, Ilya and Stengers, Isabelle, *Order Out of Chaos, Man's new dialogue with nature* (Bantam Books, 1984), pp. 27–56, 79–102, 291–313.

* * *

INTENTIONALITY AND *IRREALIA*

Burnham Terrell
University of Minnesota, Minneapolis

This paper is directed to a problem that emerges from Brentano's later writings. It starts from a set of assumptions that might rightly be taken to define a "Late-Brentano-style" approach to the philosophy of mind. These include (1) intentionality; (2) *Reismus*; (3) *Evidenz*; and (4) inner perception.[1] The problem was posed by Professor Roderick Chisholm in his article on "Intentionality" in the *Encyclopedia of Philosophy*:[2]

> . . . Brentano abandoned the ontological part of his doctrine of intentionality. In his later writings, he said that "unicorn" in the sentence "John is thinking of a unicorn" has no referential function . . . "Unicorn," in such sentences, is used syncategorematically to contribute to the description of the person who is said to have a unicorn as the object of his thought. But this conclusion seems to leave us with our problem. The statement "John is thinking of a unicorn" does not describe John as a unicorn; how, then, does "unicorn" serve to contribute to his description?
>
> The ontological problem, therefore, may be said to survive in the question, "How are we using 'unicorn' in 'John believes that there are unicorns'?" There is a temptation to say that the use of "unicorn" in such sentences has no connection at all with the use it would have in "There are unicorns." That this would be false, however, may be seen by noting that "John believes that there are unicorns" and "All of John's beliefs are true" together imply "There are unicorns."

It is true that when a word is used syncategorematically (synsemantically) it is not being used as it is when independently (autosemantically) referring. We are free, however, to draw distinctions between synsemantical expressions or between one sort of synsemantical function and another. Chisholm's question calls for such a distinction. Some synsemantical expressions are only accidentally so, by virtue of context; others are essentially synsemantic and have no autosemantic use whatsoever.

An essentially synsemantical expression could be identified with what others would call an incomplete symbol. 'Someone thinking of an object as such,' is not about what it appears to be about. It can be paraphrased, as: 'someone thinking of someone thinking of an object,' where 'object' could be said to have its ordinary role, but in a context in which it is, *per accidens*, synsemantical. Another paraphrase captures what is meant more naturally: 'someone thinking indirectly of an object.' There is no real difference in meaning between the two alternatives. To think of something indirectly is to think of it as what someone is thinking of, and that requires just that we think of someone as thinking of it.

We use words to describe mental phenomena. We use words to express mental phenomena. When John, the ill-informed zoologist, expresses his belief in unicorns by saying "There are unicorns", he uses the word 'unicorns' to specify the object of his belief. His judgment can be the primary object of a psychologist's description of John as a believer in unicorns. The psychologist uses 'unicorns' to specify the object of the belief he describes, to express a reference to unicorns incidental to a reference to someone referring to them.

Chisholm is mystified by the possibility of using a word both in referring to an object and in referring to someone who is referring to that object. We can make our own mental phenomena objects of primary reference. I can respond to a question such as "What were you referring to just now?" either by repeating the object-specifying word, 'unicorns', or by saying, "I was referring to unicorns." Someone else might respond for me, saying, "Terrell was referring to unicorns." As psychological reports, they are all equivalent. They all report the same psychological phenomenon in different ways.

The same point can be made by citing anaphoric reference in intentional contexts. If John says, "There are leprechauns," he will presumably be understood to be referring to leprechauns. Suppose that I respond, "Yes, you're right; I believe in them, too!" Chisholm's reading of Brentano would have us understand the pronoun 'them' to be synsemantical and hence disconnected from John's use of 'leprechauns.' (He would so understand the repetition of that word in the intentional context and, *a fortiori*, the use of the pronoun in its place.) I could say instead, "I believe as you do." I could say of Jerry that he believes as John and I do. Anaphoric reference and specification by similarity to a known model make no sense on Chisholm's reading. That is because his reading is mistaken.

The mistake arises from the confusion of indirect reference with synsemantical function. A description of a person as mentally active must be complex. It must include a specification of the mental act's quality and a specification of its object. A mental phenomenon is both qualitatively and objectually determined. The specifications are expressed by words in a language.

In expressing a judgment about unicorns, we use the word 'unicorn(s)' together with a component expressing the judgment's quality, in Brentano's sense. To express the objectual aspect of a description of a judgment, we use the same word, binding it to a larger context such as 'thinking about'. It is used in one case to express objectual specification for a judgment as to its primary object. In the other, the primary object of the judgment expressed by the words used is the person making the former judgment. The reference to unicorns is indirect. It provides the answer to the question, *What is his judgment about?* Any word that can be used autosemantically can be used in the expression of an indirect reference, as part of a psychological description. It functions there not independently but as an essential constituent of the specification of the mental phenomenon, the primary object of the description. Such synsemantical usage *per accidens* is not synsemantical function in Brentano's sense.

To show that the difficulties Chisholm raised are ill-founded does not satisfy his demand for a positive account of the connection between expressions of judgment and descriptions of judgers. We need to replace the ontological theory of intentional reference with a gnoseological theory in keeping with the later *Reismus*. We need a theory that will account for the logical connections Chisholm rightly insists upon without the ontological baggage of *irrealia*.

It will be useful to indroduce a distinction beween a *mental phenomenon* and an *intentional reference*. I shall call the mental phenomenon a mental act. For intentional reference I shall use that expression or 'mental reference'. An act of judgment is a complex of intentional references. It includes at least two judgmental references. One of them is objectually specified according to what we usually consider *the* object of judgment. Someone who expresses belief in unicorns by saying, "There are unicorns," is using the word to specify the primary object of his judgment, not in describing it, but in giving it verbal expression.

The act of judgment, however, includes another judgmental reference, the secondary reference of inner perception, which directly refers to the act of judgment itself and hence indirectly to the object of the primary judgmental reference. It is an affirmative judgmental reference within an act of judgment which is its object, but it is not a separate act of judgment. If it were, the first act of judgment would be its primary object and on Brentano's account it would be an act of inner observation, which he denies, not of inner perception.

To deny inner observation is not generally to deny mental phenomena that have other mental phenomena as their primary objects. There are psychologists. They make judgments about mental acts. The philosopher of mind, in making judgments about what psychologists do, must acknowledge:

a) the primary judgmental reference of some act of judgment;
b) the secondary judgmental reference, a constituent of the same act as (a);
c) the description (or affirmation) of the act of which (a) and (b) are both constituents.

The object of the secondary reference of inner perception is the entire act of which it is a constituent. A psychological judgment has a mental phenomenon as object of its primary

reference. It is distinct from its object. Nevertheless, its object, a mental phenomenon, includes as secondary reference an awareness of itself. Both (b), the inner perception and (c), the psychological judgment, are objectually specified in the same way. The secondary object of the mental phenomenon is the same as the psychologist's primary object. Therein lies the connection.

If John believes in unicorns and Jerry believes of John that he is a believer in unicorns, their respective judgments could be expressed by saying, "There are unicorns" and "John believes in unicorns." The sentences express the primary intentional references of the two judgments. John's inner perception, however, is an awareness of himself as a believer in unicorns, an affirmative secondary judgmental reference. Jerry's primary judgmental reference is an affirmation of John as believer in unicorns. Jerry judges as John does, the one in primary reference, the other in secondary reference.

Now we have a solution to Chisholm's problem. The connection between a judgment and a psychologist's account of it is provided not by introducing a common irreal object, but by taking account of the objectual specification common to the subject's inner perception and the psychologist's primary reference. It is a matter of gnoseological description. There is no need for the proliferation of ontological links to mediate between them.

ENDNOTES

[1] See the citations in my article, "Brentano's Philosophy of Mind," in: Guttorm Floistad (ed.), *Contemporary Philosophy. A New Survey* (The Hague/Boston/London 1983), Vol. 4, pp. 223–247.
[2] Edwards, Paul (ed.), *Encyclopedia of Philosophy* (New York 1967), Vol. 4, p. 202.

* * *

INTENTIONALITÄT IN TECHNOLOGIE UND TELEOLOGISCHER ERKLÄRUNG

Karl Weichselbaumer
Universität Linz

Propositionale Attitüden unterliegen genauso dem Ockhamschen Rasiermesserprinzip. Auf Quine,[1] Scheffler[2] und Stegmüller[3] zurückgehend, ist das allgemeine quantorenlogische Schema für teleologische Erklärungen daher so zu konstruieren, daß zunächst Absichten (bzw. Wünsche) sowie Glaubensinhalte (bzw. Überzeugungen) als propositionale Attitüden an Inschriften festgemacht werden. Die dadurch entstehenden inneren Sätze, die vorerst bezeichnungsundurchsichtig[4] sind, erhalten für ihre Termini eine klare ontologische Interpretation durch Reglementierung der Reichweite außenstehender Quantoren;[5] und zwar für die Hervorhebung bezeichnender Positionen aus der vorerst undurchsichtigen Konstruktion mittels Index „$_v$“, Trennstrich „ ,“ und Querverweis. Vom Gesamtsatz gesehen, enthalten dann die inneren Sätze grundsätzlich nicht-bezeichnende Terme. Zwecks leichterer Lesbarkeit, aus gleichem Grund besteht „ ,“, sind gewisse Individuenvariablen vorangestellt. (Sie könnten genauso gut unmittelbar hinter dem Prädikatbuchstaben stehen, da die Prädikatstellen nicht beliebig vertauscht werden dürfen.) Die Formel lautet dann:

$$\wedge x \wedge y \wedge z \wedge u\,[x\ W_{wv}\ x,\ y \wedge D(\vee v(Mxv \wedge Hv))y \wedge x\ G_{wv}\ x,\ z \wedge$$
$$\wedge D(\vee u \vee v(Mxu \wedge Fu \rightarrow Mxv \wedge Hv))\ z \rightarrow Mxu \wedge Fu]$$

Die Wertbereiche der Individuenvariablen sind folgend definiert: „x“ steht für Personen, „y“ und „z“ für Inschriften, „u“ und „v“ für Ereignisse. „D(−)“ ist ein prädikatbildender Operator mit Satzargument. Das Satzargument ist eine Paraphrasierung der Inschrift, für die sich nach Scheffler genau die Wahrheitsbedingungen angeben lassen. Um die logische Struktur des Argumentationszusammenhanges zu zeigen, habe ich bereits die Formel für diese sogenannten inneren Sätze eingesetzt (charakterisiert durch Atomsätze). Stegmüller, der die Grundzüge eines teleologischen Arguments als erster quantorenlogisch ausformuliert, bleibt dagegen noch bei besonderen Wünschen und steigt noch nicht zum Schema für generelle Sätze auf. Den Argumentationszusammenhang kann er daher durch eine einfachere Konstruktion herstellen. An dieser Stelle ist zu bemerken, daß die Prädikatbuchstaben nicht als Variable zu deuten sind, sondern als schematische Buchstaben, in die erst die Buchstaben als abgekürzte Prädikate einzusetzen sind. Wir würden sonst wieder in einen problematischen Platonismus geraten, den wir soweit wie möglich vermeiden wollen.

Der Ausdruck in der eckigen Klammer beginnt demgemäß in umgangssprachlicher Deutung mit „x wünscht-als-wahr von x ein y, welches eine Daß-es-gibt-ein-v-derart-,-daß-x-macht-v-,-welches-ein-H-ist-Inschrift ist“. „x“ ist als bezeichnende Position des inneren Satzes mithilfe des eigentümlichen Prädikats „W_{wv}“, das als „wünscht-als-wahr von . . .“ zu deuten ist, herausgehoben, besser: gekennzeichnet, und durch einen außenstehenden Quantor gebunden. Für „G_{wv}“, das an der Stelle der Glaubensrelation steht, und das ebenso als mehrstelliges Prädikat eingeführt ist, bei der Inschriften eine Prädikatstelle einnehmen, gilt analoges. Das Schema als ganzes gilt für die allgemeine Gesetzesaussage im teleologischen Argument, das im Syllogismus des modus ponens „$\forall u \forall x(Mxu \wedge Fu)$“ durch Deduktion liefert und das Explanadum „$Mab \wedge Fb$“ daher befriedigt.

Aristoteles kommt zu einer Systematik der Ursachen, der seine eigentümliche Weltsicht, der des bildenden Künstlers,[6] zugrunde liegt. Die Zweckursache (zusammen mit der Formursache) nimmt in diesem Fachwerk eine ganz andere Stellung ein, als sie in der heutigen Hand-

lungstheorie einnimmt. Das Telos hat nicht bereits mit dem Anstoß zur Handlung seine Schuldigkeit getan. Der Schwerpunkt teleologischer Erklärung liegt nach Aristoteles daher nicht wie bislang beim Anstoß der Handlung, er liegt bei der Gestaltung der Handlung selber. Für ihn ist sogar die Gestaltung allein das Explanandum.

Selbst Wright, der als führender Intentionalist gilt, entzieht sich der herkömmlichen Sichtweise nicht. Er konzentriert sich immerhin auf die logische Grammatik intentionalen Sprachgebrauchs insoweit, als sich psychische Phänomene intentionaler Art dort sedimentiert haben.[7] Er folgt dem Leitgedanken: „Wir sprechen, machen Äußerungen, und erst s p ä t e r erhalten wir ein Bild von ihrem Leben" (*PU* XI) – prima facie das vernachlässigend, wieweit im jeweiligen Fall die Ausdrücke für die Realität zutreffen. Hier werden aber die intentionalen Phänomene nur indirekt, nicht sie selber behandelt.

Freilich sind die menschlichen Aktivitäten der Aristotelischen teleologischen Erklärung in verschiedenem Maße zugänglich. Das Formen des Tons steht am einen Ende der Skala, der Hausbau in der Mitte, das Ziehen am Flaschenzug am anderen. Kunst und Technik gehen fließend ineinander über. Wie bereits angedeutet, hat Aristoteles mit der Endursache nicht den Beginn der Aktivität im Auge, wenn er mit ihr etwas erklärt, sondern das Resultat, etwa eine Reiterfigur. Die Vorstellung der Form in der Seele des Baumeisters sei Bewegursache, vergleichbar dem Samen, der in die Erde als Material der Form gesenkt wird. Die Formen als solche aber sind zeitlos, und neben die Formursache stellt sich die Zweckursache als treibende Kraft, die mit ersterer zusammen auftritt. Form- und Endursache erklären die fertigen Dinge in ihrer Beschaffenheit und ihren Weg dorthin, nicht aber den Beginn der Materialformung. Dies leistet bereits die Bewegursache. (Sie darf nicht mit der Stoffursache verwechselt werden, die zugleich als Störursache Widerstand leistet und für etwaiges Mißlingen verantwortlich ist. Die List der Vernunft, mit Hegel zu sprechen, hat sie in Gestalt der Technik großenteils gezähmt.)

Versuchen wir, die teleologische Erklärung im Aristotelischen Sinn zu fassen,[8] wenn auch modernisiert, so gelangt man zu folgendem Schema für teleologische Erklärungen:

$$\begin{aligned}&\wedge x \wedge y \wedge z \wedge u \wedge v\,[x\ W_{wv}\ x,\ y \wedge D(\forall v(Mxv \wedge Hv))\ y \wedge x\ G_{wv}\ xyz,\ z \wedge\\ &\wedge D(\forall u \forall v(yz \Box Mxu \wedge Fu \rightarrow Mxv \wedge Hv))\ z \rightarrow\\ &\rightarrow (yz \Box Mxu \wedge Fu \rightarrow Mxv \wedge Hv)].\end{aligned}$$

Es gilt wieder für die allgemeine Gesetzesaussage im Argument, überraschenderweise aber auch für die von Technologien, letztlich für Technologie im strengen Sinn selber, die sich dann als teleologische Prognose erweisen, wenn die Spezialisierung im Syllogismus reziprok (analog zum Hempel-Openheimschen Erklärungs-Prognose-Schema) erfolgt. Explanandum und zu Prognostizierendes ist dann durch „$\forall x \forall v(Mxv \wedge Hv)$" erfüllt. Daß das Objekt der Erklärung sich immer auf einen Aspekt reduziert, präzise: auf singuläre Satze,[9] spricht für das Bestehen eines bislang unbekannten Aspekts, nämlich des gezeigten Aristotelischen. Er ist sogar eigentliches wie zentrales Objekt der TE.

Dieser Aristotelische Aspekt läßt mit Putnam überdies eine neue Deutung zu: Der von mir vorgeschlagene Vierstellen-Prädikate bildende Operator „$\Box$" mit dem Zweistellen-Prädikat „M" zum Argument, läßt sich in obiger Formel als Intentionalitätsoperator deuten. Er überführt – unbeschadet extensionaler Deutbarkeit – ein Machen „M" in ein intentional durchdrungenes Machen „M′". Lediglich leichterer Lesbarkeit wegen wiederum werden „y" und „z", die Inschriften, die zusätzliche Prädikatstellen belegen, vorangestellt. Putnam legt auf Fields Tarski-Kritik anhand ihrer Unzulänglichkeiten die Grenzen der naturwissenschaftlichen Methode im impliziten Wissen dar. Es ist in Theorie nicht ausformulierbar, da es sich in Fertigkeit gründet.[10] Die generelle Formel für die Gesetzesaussage im teleologischen Argument bildet bei der teleologischen Prognose, die zugleich technologische Prognose ist, die Technologie im strengen Sinn (das Können, das das Sollen impliziert). Gerade durch den eingeführten Intentionalitätsoperator wird dieses implizite Wissen berücksichtigt, das zum expliziten hinzutreten muß, um nach Putnam den Praxisvollzug – und damit auch von Technik –

überhaupt zu gewährleisten. Die Einbeziehung erfolgt aber nur indirekt über die Intentionalität (in Verbindung mit den propositionalen Attitüden), die die Aktualisierung des impliziten Wissens sicherstellt, und zwar material; ohne im Satz formal ableitbar zu sein. Der Operator steht gleichsam für eine Interpolation der exakten Teilkurve zum Kurvenstützpunkt „Absicht", wobei dessen Lage so angegeben ist, daß die Interpolation gelingt. Erklären i.e.S., aber auch wie bei Putnam Verstehen,[11] werden dann ebenso durch diesen Operator differenziert und in ihrem Standort erhellt.

Diese Strukturierung von Technologie mit unentbehrlicher Supplementierungsfunktion von Theorie erlaubt einen – de facto von vornherein vollzogenen – zwanglosen Übergang zur praktischen Zweck-Mittel-Relation, deren Eigentümlichkeiten aber auf der hier erarbeiteten Folie sich erst abheben.[12]

ANMERKUNGEN

1 Quine, W. v. O., *Wort und Gegenstand (Word and Object)* (Stuttgart 1980).
2 Scheffler, I., *The Anatomy of Inquiry: Philosophical Studies in the Theory of Science* (New York 1963). Vgl. dazu W. v. O. Quine, *Ontological Relativity and Other Essays*, deutsch: *Ontologische Relativität und andere Schriften* (Stuttgart 1975).
3 Stegmüller, W., *Probleme und Resultate der Wissenschaftstheorie und Analytische Philosophie* Bd. I (Berlin-Heidelberg-New York 1974).
4 Whitehead, A. N. und Russell, B., *Prinicipia Mathematica* Bd. 1 (Cambridghe ²1925).
5 Quine, W. v. O., *Wort und Gegenstand*, aaO., bes. S. 289ff.
6 Windelband, W., *Lehrbuch der Geschichte der Philosophie*, H. Heimsoeth (Hrsg.) (Tübingen ¹⁵1957). Das Weltgeschehen wird „in letzter Instanz unter der Analogie des bildenden Künstlers betrachtet, der für die Verwirklichung seines gestaltenden Gedankens in dem spröden Material eine Grenze findet" (ebd., S. 123).
7 Wright, G. H. v., *Explanation and Understanding* (London 1971).
8 Siehe hiezu Sachsses Plädoyer für frühere Stadien der Kategorienbildung. (H. Sachsse, *Kausalität-Gesetzlichkeit-Wahrscheinlichkeit* (Darmstadt 1979).)
9 Stegmüller, W., aaO.
10 Putnam, H., Meaning and the Moral Sciences (London 1978).
11 Ebd.
12 Ich verweise insbesondere auf Max Webers „formale Zweckrationalität" und die nun bis Putnam reichende Diskussion (H. Putnam, *Reason, Truth and History*, deutsch: *Vernunft, Wahrheit und Geschichte* (Frankfurt a. M. 1982).)

* * *

INTENTIONS AS *INTENTIONES ANIMAE* IN OCKHAM AND AS IMMANENT MENTAL 'OBJECTS' IN BRENTANO

Elisabeth Leinfellner-Rupertsberger
Lincoln, Nebraska

This paper focuses on three topics: (1) two paradigms of semantics and an attempt to unite them by using pragmatic and psychological concepts; (2) a characterization of Ockham's discussion of *intentio* in the framework of these two paradigms and as their synthesis; (3) an interpretation of some fragments from Brentano's writings in the light of (1) and (2).

The point of departure for relating a Scholastic philosopher to Brentano is the well-known quote from Brentano's *Psychology from an Empirical Standpoint*:

> Every mental phenomenon is characterised by what the Scholastics of the Middle Ages called the intentional (or mental) inexistence of an object, and what we might call, though not wholly unambiguously, reference to a content, direction toward an object (which is not to be understood here as meaning a thing), or immanent objectivity.[1]

Spiegelberg and Marras have examined Brentano's relationship to Scholastic philosophy from a mainly Thomistic standpoint.[2] This is justified considering that, f.i., *Psychology from an Empirical Standpoint* makes reference to Ockham but once, when Brentano discusses the question of whether the existence of a being is the same or a different reality than the being itself. According to Brentano, Ockham is correct in denying the difference, but arrives at the wrong conclusions.[3] On the other hand, Aquinas—and Aristotle—figure much more prominently in *Psychology from an Empirical Standpoint*.

However, the fragments from Brentano which I am going to discuss and which fall into a much later period than the (German) 1874 edition of *Psychology from an Empirical Standpoint* come much closer to Nominalism and to Ockham in particular than is to be expected from someone who had earlier (1894) relegated Ockham's philosophy to the second stage of medieval philosophy's decline, i.e., scepticism. In 1901, Brentano went as far as admitting —with reservations—that Nominalism might be right at least on some accounts. After discussing grammatical abstracta and universals in a way outlined below, he continues: "A concession should be made to Nominalism; not that Nominalism is correct, however, but only that certain errors can be purged from the opposing theory."[4]

I do not want to claim that Ockham's philosophy had any influence on Brentano's. The similarities between these two philosophies show, however, that similar problems may lead to similar solutions, a phenomenon upon which we sometimes bestow the grandiose name of *philosophia perennis*.

Spiegelberg points out that the above quote contains two characterizations of mental phenomena:[5] Of one, intentional inexistence, he claims that it is indeed a Scholastic, particularly Thomistic, concept; of the other, reference to a mental object, that it originated with Brentano.[6] According to Spiegelberg, there is in Scholastic philosophy no significant connection between reference to a mental object and *intentio*.[7] Marras, on the other hand, attempts to show that in Aquinas *intentio* is by nature representational and therefore has to stand in a reference relation to an object.[8] We have here two interrelated views of referential semantic relationships. In Brentano, the German term of which 'reference' is a translation means a relation (something relation-like, according to Brentano) between someone who thinks and the immanent mental object.[9] In modern philosophy, 'reference' means a relation between a linguistic expression and some object which may be real, but also possible. In Ockham, '*suppositio*' overlaps with 'reference' in the modern sense; but Ockham also discusses *intentio* in connection with supposition (see below).

However, it seems that Marras stresses the empirical object as referent, while what Spiegelberg has in mind is reference in the sense of Brentano as reference to an immanent, intentional object. Marras is correct in assuming that *intentio* in Aquinas is representational of an object and therefore intrinsically referring; but this does not solve the problem of whether the Scholastics discussed reference to an intentional object.

Ockham's concepts of a mental language and of *supposito simplex* furnish arguments against Spiegelberg's contention that Scholastic philosophy does not deal with reference to intentional objects; in addition, they allow us to broaden Marras' discussion of referring *intentiones* to the discussion of reference to *intentiones*.

Ockham distinguishes between three kinds of expressions: spoken, written, and mental. The spoken and the written language originate by convention (*thesei*), mental language by nature (*physei*). Ockham identifies the mental terms with the *intentiones animae*; they may be universals or singular terms.[10] The written language is an extension of the spoken language, the spoken language an extension of the mental one. However, a written term signifies independently from a spoken one, and a spoken one independently from a mental term.[11]

Suppositio simplex obtains when a term signifies secondarily and stands for a mental term, that is, an *intentio*, as *suppositum*.[12] Since supposition involves reference it is obviously possible to speak of reference to an intentional object, the mental term, in Ockham's philosophy. And since the *supponens* of *suppositio simplex* may be a mental term, too, we may be faced with a totally intentional and at the same time referential relation.

We have just used the term 'signifies secondarily'; Ockham explains the difference between primary and secondary signification as follows: In contrast to connotative terms, an absolute term like 'whiteness' has in *suppositio personalis* only primary signification: it denotes something in the empirical world, and nothing else besides; 'white', on the other hand, has in *suppositio personalis* both primary and secondary signification. When 'white' has secondary signification, it connotes 'whiteness'.[13] Since, as Ockham says, whiteness is not white, this secondary relation of signification is a language-immanent semantic relation (pardigm (2); see below). As to be expected, there exists a nominal definition where 'white' is defined in terms of 'whiteness'; 'whiteness' however is an absolute term and cannot really be defined nominally.

Terms in *suppositio personalis* signify primarily, and some, moreover, secondarily; terms in *suppositio simplex*—and *suppositio materialis*—signify only secondarily. One can see that Ockham treats relations between terms (*suppositio materialis* and *simplex*) totally different from relations between terms and empirical objects (*suppositio personalis*); this includes the interesting case of intentional objects as referents and elements of a mental language (*suppositio simplex*). Brentano stresses the same point when he comments that not each intentional object has a physical counterpart: mental reference is, therefore, not a genuine relation but something relation-like.[14]

In the spirit of *philosophia perennis* (see above), I will now relate Ockham's and Brentano's views on intentions, reference, etc. to two paradigms of semantics and a synthesis of these two paradigms, 'paradigm' taken in its everyday sense.[15]

Paradigm (1) we may call the truth-functional or referential paradigm; semantics, then, is of a Frege, Montague, or Davidson type. If we stick here to a strictly empiristic position, that is, want to avoid platonistic referents or referents in possible worlds, we must admit the existence of a large class of expressions which have no meaning at all, f.i. fictitious and metaphysical terms. A way of avoiding platonistic referents or referents in possible worlds would be to introduce mental or psychological intentional objects or intentions. However, in the last decades, and particularly since Carnap's *Logischer Aufbau der Welt* of 1928, empiristic philosophers have shunned intentional objects or intentions, since they fell under the dogma: "Thou shalt not commit the sin of psychologism". Ockham has been criticized, too, for having committed psychologism,[16] and so has Brentano.[17]

However, from the linguist's standpoint the introduction of mental or psychological intentional objects or intentions alone is not sufficient if we want to develop a comprehensive

semantic system, expecially if we consider Brentano's argument that there is no one-one correspondence between thought and language.[18]

Paradigm (2) may be called the 'structural' or 'contextual paradigm'; it does away with the notion that reference and truth functions are the basis of semantics. Instead, we recur to language use f.i. in the sense of the later Wittgenstein and define meaning (or, rather, to have meaning) as having statistically invariant use in a language. Language use is identified with the set of all statistically invariant contexts of a given term. The linguist refines the set of statistically invariant contexts by introducing, f.i., systems of semantic markers, semantic networks, and the like.[19] This is the theoretical expression of Humboldt's famous dictum that one single word evokes the entire language. For sentences we make use of semantic filters, selectional restrictions, semanticized case grammars, and the like.

Paradigm (2) comprises componential analysis, many systems of Artificial Intelligence and Knowledge Representation, semantic networks, memory models, and the like. Its disadvantage is, of course, that all relations to the empirical world may disappear; we would end up with a truncated semantics.

Paradigm (2) can be seen as an attempt to avoid psychologism,[20] as evidenced by § 501 of *Philosophical Investigations* where Wittgenstein ironically comments on the notion that each sentence expresses a thought. Which thought is expressed by the sentence "It rains"?, he asks.

Even from this shortest of sketches it is obvious that a semanticist attempting to comprehensively explain semantic phenomena from an empiristic standpoint somehow has to use and unite the two paradigms. One way of doing this is by introducing a thesis which we may call 'invariance thesis'. This thesis states that statistically invariant empirical phenomena lead to statistically invariant contexts, that is, meaning as statistically invariant language use.

The most likely candidate for the missing link between the two paradigms is an intentional, psychological, or pragmatic connection. I plead here for memory, particularly also in the light of the most recent developments in memory research, where one now distinguishes between (at least) two kinds of 'linguistic' memory, episodic and conceptual ('semantic'). Episodic memory corresponds to paradigm (1), conceptual memory to paradigm (2). Likewise, in Artificial Intelligence one works with two components of semantic networks.[21]

How is such a two-levelled semantics and its linking component, the intentional, psychological (mental), or pragmatic element, anticipated in Ockham's system?

In Ockham's writings, *suppositio personalis* clearly corresponds to referential, truth-functional meaning according to paradigm (1). A term exhibits *suppositio personalis* if it stands for what it primarily signifies (see above).

Paradigm (2) appears in Ockham's philosophy in the following form: Paradigm-(2) meaning is represented by *suppositio materialis* as described under (b). Ockham assumes that there are two kinds of suppositio materialis: (a) *suppositio materialis* is a meta-relation as in "'woman' is a noun" where 'woman' has *suppositio materialis*; (b) *suppositio materialis* in connection with the problem of figment terms like 'chimera' and 'stag-goat'. If the term 'chimera' in "A chimera is an animal that is composed of a goat and a cow" is interpreted as having *suppositio personalis*, the sentence is false. If we change *suppositio personalis* to *suppositio materialis* by substituting 'signifies (secondarily)' for 'is', we arrive at the nominal definition, "'Chimera' and 'animal composed of a goat and a cow' signify the same", which is true.[22] According to some interpreters of Ockham, this process can be repeated iteratively until we arrive at terms capable of *suppositio personalis*,[23] that is, at paradigm (1).

This brings us to the question of how Ockham mediates between the two paradigms. On my interpretation, it is the universal (the universal mental term) in connection with *suppositio simplex* which mediates.

Let us take as an example universals as mental terms which are grammatical concreta and which belong to the category of substance, f.i. the term 'man'. Such a universal (universal mental term) is absolute, that is, does not connote and—in *suppositio personalis*—does not signify secondarily. In predicate position, it does not denote a property, nor does it denote a platonistic or *in-re* universal. In the sentence "Socrates is a man" both 'Socrates' and 'man'

personally supposit for the same empirical entity in which case the sentence is true. Such a universal is a mental term which is capable of standing for many things, one by one, so to speak, things which are similar to one another. This process of "standing for" is, however, not an abstraction, in the same way as a paradigm-(2) meaning is not the result of an abstraction. Thus an Ockhamist universal does not denote the class as a whole, nor does it denote a class property. The universal collects, as it were, the various referents and in this respect corresponds to the set of statistically invariant contexts since it is at the same time a mental term. Moreover, the term which is said to have meaning with respect to the universal as *intentio animae* is the one which stands in *suppositio simplex* to it.

It is interesting to note that Ockham interprets the universal mental term, a term which stands for many things, at the same time as an act of understanding. This stresses the intentional character of the universal. It is not surprising, then, that in the appendix (1911) to *Psychology from an Empirical Standpoint* Brentano replaced the concept of an immanent mental object with the concept of mental activity, or, rather, the mentally active subject.[24]

We come now to the comparison between elements of Ockham's and Brentano's philosophy as regards the universal and related matters. Brentano's views are expressed primarily in *Kategorienlehre* (1933, 1968) and in his letters (from *Wahrheit und Evidenz* (1930)).

The first point of our comparison concerns the doctrine that the universal is not a thing outside the mind. Ockham concedes that there might be conventional—as opposed to natural—signs which are universals, that is, spoken or written universals. But he states that he has no interest in them. He devotes an entire chapter of the *Summa* to the demonstration that universals are natural signs in the mind, *intentiones animae*.[25]

Brentano holds a similar view. In a letter to Marty of March 1901 he says: "There is nothing universal in the things; the so-called universal, as such, is only in the one who is thinking".[26] In a passage from *Kategorienlehre* (dated 1916) he maintains, like Ockham, that there are universal as well as particular intuitions (*Vorstellen*). However, in an appendix to the 1924 edition of *Psychologie vom empirischen Standpunkt* (dictated 1917) he presents a view even more restricted than Ockham's: Even if we think intuitively (*anschaulich*), "we can still think only in universal terms". The universal, says Brentano, has, however, only a secondary existence in the mind, not the primary existence of an empirically existing thing.[27]

This secondary existence of the universal according to Brentano is obviously systematically related to Ockham's notion that a term in *suppositio simplex* (and *suppositio materialis*) has only secondary signification.

The second point of our comparison between Ockham and Brentano is the relation of the universal to the empirical world. As we have already noted, in Ockham the universal mental term is characterized by being able to stand for many things, one by one. Brentano distinguishes between three approaches to the problem of universals, the first two of which he considers to be wrong:[28]

(1) The universal exists empirically. This is wrong, since nothing indeterminate (*unbestimmt*)—which Brentano identifies with the universal—can exist. (2) There is nothing in the empirical world which corresponds to the universal. This is wrong, too, since (3) the universal is characterized exactly by the fact that many empirical objects, phenomena may correspond to it in the same way.[29]

The third point of comparison is the problem of whether a universal corresponds to (Brentano), or denotes (Ockham), a part of a substance, or not. I will start here with Brentano's view as expressed in the letter to Marty of March 1901. Brentano at first distinguishes between grammatical concreta and grammatical abstracta and then repudiates his earlier view according to which a term like 'redness' denotes a part of a thing, in analogy to things having genuine parts. Brentano now offers another solution: When the term 'redness'—which is the linguistic counterpart of a mental universal—is applied to a red object, the whole object is thought of by means of an indefinite concept, the mental universal itself.[30]

Ockham's views, as far as they can be gleaned from his writings, are very similar. At first, like Brentano in *Kategorienlehre*, he denies that grammatical concreta and grammatical

abstracta are synonymous, even though Aristotle maintains that they are. Thus 'whiteness' and 'white' are not synonymous.[31] Ockham places 'whiteness' directly into the category of quality, and 'white' "by reduction".[32] In *suppositio personalis*, 'whiteness' signifies only primarily and is, therefore, an absolute term, 'white', however, is absolute as well as connotative; in the latter case 'white' secondarily signifies 'whiteness'.

In *SL* II, 2, which deals with the truth of singular, non-modal propositions, Ockham states the principle that a sentence is true only if both subject and predicate supposit for the same. This neither requires that subject and predicate are identical, nor that, f.i., the term 'humanity', a grammatical abstract from the category of substance, supposits for a part of, e.g., Socrates.[33] The universal term 'whiteness', a grammatical abstract from the category of quality, poses different, and more confusing, problems. Ockham's viewpoint has been interpreted as indicating that the universal 'whiteness' primarily signifies not one but many whitenesses, and exactly as many whitenesses as there are white things.[34] This interpretation runs into the same difficulties as the view to which Ockham explicitly objects, namely, that there is a universal object whiteness. In the framework of Ockham's theory of truth both views lead to an unexpected consequence: A sentence in which 'whiteness' is predicated of a term from the category of substance can never be true, since the condition of "suppositing for the same" can never be fulfilled. We might want to return to the interpretation that 'whiteness' supposits for white things. But in *SL* I, 33 Ockham does say that 'whiteness' primarily signifies whiteness, and in *SL* I, 5 that 'white' and 'whiteness' do not primarily signify the same. Implicitly, this amounts to a prohibition of sentences containing abstracta from the category of quality as predicates in combination with a term from the category of substance. Indeed, such sentences are conspicuously absent from Ockham's example sentences.

At any rate, Ockham's universal term 'whiteness' is just as indefinite as Brentano's 'redness' and its corresponding mental universal redness, since both terms are applied to many empirical phenomena.

This is a typical instance of "reism" in Brentano[35] as well as in Ockham; according to this, only individual things exist in the sense proper.

We come now to the fourth and last point of comparison between Ockham and Brentano, the nature of grammatical abstracta. Brentano concludes in the 1901 letter to Marty that grammatical abstracta are delusions, useful ways of speaking.[36] This implies that they somehow have to be reduced to their respective grammatical concreta—at least from the epistemological standpoint.

It is interesting that Ockham has introduced a kind of reductive analysis for grammatical abstracta from the category of substance.[37] While absolute terms from the category of substance which are grammatical concreta, like 'man', 'animal', pose no particular problem, their respective grammatical abstracta, 'humanity', 'animality', do, since one cannot say, f.i., "Animality runs" in the same way as one says "An animal runs".

In *SL* I, 8 Ockham applies a reductive analysis to terms like 'humanity'. The above example shows that the elements of the pair 'man'-'humanity'—and similar pairs—are not synonymous. On the other hand, both terms may primarily signify, and personally supposit for, man. This is incongruous. Ockham solves the puzzle in the following manner: 'humanity' is equivalent in primary signification to a combination of terms. One element of this combination is a concrete term and we may assume that it is this concrete term which is responsible for the grammatical abstract's primary signification and personal supposition. The other element of the combination consists of one or more terms which, according to Ockham, are syncategorematic. Thus, 'humanity' primarily signifies the same as 'man as man' or 'man necessarily . . .', 'as man' and 'necessarily . . .' being the syncategorematic elements. Compared to the discussion of 'syncategorematic' in *SL* I, 4, we see that here a much broader concept of syncategorematic has been employed. At any rate, 'syncategorematic' can be viewed as a technical term from a paradigm-(2) semantics.

The syncategorematic elements 'as man' and 'necessarily . . .' attempt to express something theoretical about man.[38] Thus 'humanity' and similar terms are related to theoretical con-

structs which, on the one side, may be seen as purely theoretical (paradigm (2)) and, on the other side, may be joined to the empirical world (paradigm (1)).

In order to maintain the possibility of personal supposition for 'humanity', 'humanity' has been epistemologically reduced to 'man'; and that aspect of grammatical abstracta which Brentano termed 'delusions' Ockham calls 'syncategorematic'.

From the discussion of Brentano's and Ockham's views in the framework of the two semantic paradigms we can now draw one main conclusion, (1), and two subordinate conclusions, (2) and (3); (2) we may label 'content', and (3) 'form'. (1) Intentionality mediates between the two paradigms of semantics, both of which are necessary for a systematic discussion of all semantic phenomena. (2) Content: The concept of intentionality has to be applied to all semantic phenomena, not only to beliefs, modalities, etc. (3) Form: In our framework, we cannot restrict the analysis of intentionality to the analysis of intentional language, to the logic of intentional expressions, etc., in short to the Analytic philosopher's point of view. This is not to say that the Analytic view is to be rejected, but only that it is too narrow for our purposes: empiristic semantics has to include psychological, biological, and brainphysiological inquiries. Strict empiricism of the Vienna-Circle type drove out psychologism, but psychologism is about to enter through the back door.

ENDNOTES

[1] Brentano, F., *Psychology from an Empirical Standpoint*, ed. L. L. McAlister, transl. A. C. Rancurello, D. B. Terrell, and L. L. McAlister (New York 1973), p. 88.

[2] Spiegelberg, H., "'Intention' and 'Intentionality' in the Scholastics, Brentano and Husserl", in L. L. McAlister (ed.), *The Philosophy of Brentano* (Atlantic Highlands, N. J., 1977); Marras, A., "Scholastic Roots of Brentano's Conception of Intentionality", in McAlister (1977).

[3] Brentano (1973), p. 229.

[4] Brentano, F., *Die vier Phasen der Philosophie und ihr augenblicklicher Stand*, ed. O. Kraus (Leipzig 1926), p. 7ff.; Brentano, F., *The True and the Evident*, ed. R. M. Chisholm, transl. R. M. Chisholm, I. Politzer, and K. R. Fischer (London-New York 1966), p. 64.

[5] Spiegelberg, H., *The Phenomenological Movement* Vol. 1 (The Hague ^{2}1969), p. 40f.

[6] Ibid.; see also Marras (1977), p. 128f.

[7] Cf. Marras (1977), p. 136.

[8] Marras (1977), p. 133ff.

[9] Brentano (1973), pp. 88, 272. For a detailed exposition of this and related questions see R. M. Chisholm, "Intentional Inexistence", in McAlister (1977).

[10] Ockham, Wilhelm von, *Summa logicae; pars prima*, ed. Ph. Boehner (Paderborn 1962), cap. 3 = *SL* I, 3. Ockham himself denies that there is a mental language: there are only mental terms and mental propositions, see *SL* I, 1.

[11] *SL* I, 1.

[12] *SL* I, 64, 68.

[13] *SL* I, 33; Ockham, Wilhelm von, *Summa logicae*: *pars secunda et tertiae prima*, ed. Ph. Boehner (Paderborn 1962), pars secunda, cap. 7, 10 = *SL* II, 7, 10; Spade, P., "Ockham's Distinction Between Absolute and Connotative Terms", *Vivarium* Vol. 13 (1975); Freddoso, A., "Ockham's Theory of Truth Conditions", in *Ockham's Theory of Propositions*, transl. A. F. Freddoso and H. Schuurman (Notre Dame, Ind., 1982), p. 4ff. According to *SL* I, 67 and 68, all terms may exhibit *suppositio materialis* and *simplex*; therefore, all terms may signify secondarily, but not in *suppositio personalis*.

[14] Brentano (1973), p. 272; see also Brentano (1966). p. 77.

[15] For a more detailed exposition of Ockham's philosophy in the framework of the two paradigms see W. Leinfellner and E. Leinfellner, *Ontologie, Systemtheorie und Semantik* (Berlin 1978), chap. 1.4; W. Leinfellner and E. Leinfellner, "Ockhams Semantik und Pragmatik" in H. Stachowiak (ed.), *Handbuch der Pragmatik* Vol. 1 (1984, in press).

[16] Hickman, L., "Three Consequences of Ockham's 'Mental-Act' Theory", *Southwest Philosophical Studies* Vol. 10 (1979).

[17] Brentano (1973), p. 306f.

[18] Brentano (1966), p. 71.

[19] One of the earliest (if not the earliest) attempts to use Wittgenstein's concept of meaning in use as philosophical background for a statistical approach to word meaning is W. Z. Shetter and F. W. Blaisdell, "Altsächsisch *mahlian* und die Verben des Sprechens und Sagens", *Zeitschrift für deutsche Wortforschung* Vol. 18 (Neue Folge Vol. 3) (1962).
[20] Cf. E. Leinfellner, "Zur nominalistischen Begründung von Linguistik und Sprachphilosophie: Fritz Mauthner und Ludwig Wittgenstein", *Studium Generale* Vol. 22 (1969), p. 216ff.
[21] Tulving, E., *Elements of Episodic Memory* (Oxford-New York 1983), p. 17ff.; Leinfellner, E., Steinacker, I., and Trost, H., "Anaphoric and Kataphoric Relations, Textual Cohesion, and Reference Resolution", *Computers and Artificial Intelligence* (1984, in press); Leinfellner, E., "Connexity/ Coherence and the Semantic Net" in M.-E. Conte, J. S. Petöfi, and E. Sözer (eds.), *Text and Discourse Connectedness* (1984, in press); Leinfellner, E., "Semantic Nets: Their Linguistic Usefulness", in P. Simons (comp.), *7th International Congress of Logic, Methodology and Philosophy of Science* (Salzburg 1983). See also the paper by Werner Leinfellner in this volume.
[22] *SL* I, 26; *SL* II, 14.
[23] F. i. Spade (1975), p. 69f.
[24] *SL* I, 15; cf. Leinfellner-Leinfellner (1984, in press), chap. 3.4; Brentano (1973), p. 276.
[25] *SL* I, 15.
[26] Brentano (1966), p. 64.
[27] Brentano, F., *Kategorienlehre*, ed. A. Kastil (Hamburg 1968), p. 21; Brentano (1973), pp. 311, 275f.
[28] Brentano (1968), p. 21.
[29] Brentano (1968), p. 21f.; Brentano (1966), pp. 72, 99.
[30] Brentano (1966), p. 63f.
[31] Brentano (1968), p. 6; *SL* I, 7.
[32] *SL* I, 41.
[33] *SL* I, 7.
[34] Loux, M. J., "The Ontology of William of Ockham", in *Ockham's Theory of Terms,* transl. M. J. Loux (Notre Dame, Indiana-London 1974), p. 13f. (misprint on p. 13, but see p. 14).
[35] See, f.i., T. Kotarbinski, "Franz Brentano as Reist", and D. B. Terrell, "Brentano's Argument for Reismus", both in McAlister (1977).
[36] Brentano (1966), p. 64.
[37] *SL* I, 7,8.
[38] Cf. Loux (1974), p. 11.

* * *

INTENTIONALITY, METHODOLOGICAL SOLIPSISM AND COMPUTATIONAL PSYCHOLOGY*

Ausonio Marras
University of Western Ontario

The development of cognitive psychology in the last twenty years has given new respectability to the study of mind. The positing of "inner" mental states and processes, once regarded as methodologically suspect, is now widely regarded as indispensible for a scientific understanding of intelligent behaviour. Controversy mainly arises about the *nature* of such inner, causally efficacious states, and about the conceptual resources which are appropriate for their description and identification.

According to Fodor (1981) and other theorists, the relevant mental states are those which essentially correspond to the *propositional attitude* states of commonsense, folk psychology. On this construal, cognitive psychology should be viewed essentially as a scientific extension of folk psychology, and thus as committed to the idea that the regularities in human behaviour can best be captured in terms of people's beliefs, wants, expectations, and other typically *intentional* states. Accordingly, the task of cognitive psychology is to provide a theoretical framework in which the loose concepts and generalizations of folk psychology can be analyzed, systematized, and scientifically tested. The scientific study of mind is thus, on this account, the study of the nature and dynamics of the propositional attitudes; it is, quite simply, the *science of the intentional.*

Since the main burden of this paper is to consider a challenge to this interpretation of the cognitivist program, it is important to consider in more detail some of its conceptual and methodological commitments.

On Fodor's account, propositional attitudes are to be viewed as relational states of organisms: they relate organisms to *mental representations*. Mental representations, like linguistic representations, are essentially *symbolic*, and have thus both *formal* and *semantic* properties. The intentionality of propositional attitudes derives from the semantic (intentional) content of the representations which are their relata; their causal role, on the other hand, is consequent upon the formal features of the representations. This, in a nutshell, is the conceptual skeleton of Fodor's "Representational Theory of the Mind" (RTM).

Because mental representations are construed as symbolic, it is natural to give the RTM a computational interpretation: the mental processes involved in cognitive activity are essentially computational in the sense that they are rule-governed operations defined over formally specifiable representations. On this construal, to be in a propositional attitude state is to bear a computational relation to a symbol-structure representing the object of the propositional attitude.

A virtue of the computational model of the mind is that, if sound, it would provide a solution to the problem of how mental states can be causally efficacious and, in particular, of how they are able to interact with one another in the production of behaviour in such a way as to satisfy the rules of theoretical and practical reasoning which apply in the intentional domain. For we know that computing machines can execute intentionally or semantically interpretable operations because those operations are formally representable, and if they are formally representable they are Turing reducible and thus mechanically realizable. The computer model thus enables us to relate the intentional domain to the physical domain, and thus provides a solution to the vexing old problem of how *reasons* can be *causes*.

It is clear that the computational form of the RTM imposes very specific constraints on the theory. Since only formally encoded aspects of a represented domain can effect a computing

machine's functioning, all semantically relevant characteristics of a represented domain must be encoded in the syntactic structure of a computational sequence representing an operation in that domain. This condition, which Fodor (1980) has called the *formality condition*, will insure that a symbol using device will be able to operate on symbol tokens in virtue exclusively of the formal-syntactic properties of those tokens, and thus that it will behave in semantically meaningful ways without having to concern itself with *meanings* but only with *forms*.

Promising as it seems, this type of cognitivist program has been challenged by some who do not believe that there can be a science of the intentional, or who do not believe that a mature scientific psychology should endeavor to accommodate the propositional attitude framework of folk psychology.

An instructive illustration of this type of challenge is Steven Stich's (1983) recent critique of Fodor's RTM. Stich distinguishes two forms of the RTM, a *strong* and a *weak* form, and charges Fodor with equivocating between the two. Both forms of the RTM share the intentionalist assumption that mental states are relations between organisms and "contentful" or semantically interpreted representations. But whereas the strong RTM insists that the nomological generalizations which detail the causal interactions among mental states apply to them by virtue of their intentional *contents*, the weak RTM denies this and holds instead that those generalizations apply to mental states exclusively in virtue of their *syntactic* properties. Against both forms or the RTM, Stich advances a strictly *syntactic* theory of mind (STM), which accepts the distinguishing claim of the weak RTM, but declares itself agnostic on the question whether mental states *have* content or other semantic properties. More precisely, whether or not the mental states posited by a mature cognitive science have semantic properties, the STM claims that there is *no need* to suppose that they do. It follows that according to the STM there is no need to construe the objects to which mental states relate us to as essentially *symbolic* or *representational*: according to the STM, the whole intentionalist framework of folk psychology which identifies mental states with propositional attitude states may well be jettisoned with no detriment to cognitive science.

There is no space, here, to assess the merits of Stich's STM, nor the soundness of his claim that Fodor has inconsistently endorsed both forms of the RTM. As a matter of fact, I believe that the STM is untenable, and that, correctly interpreted, Fodor's position commits him to the strong, not the weak, form of the RTM; but I am not going to argue that here. What I am going to argue for is that the strong RTM is immune to a criticism levelled by Stich against it. The criticism, essentially, is that the strong RTM is inconsistent with a methodological principle which is held to be central to any theory which, like the RTM, construes mental states as *computational*. This is the principle of *methodological solipsism*, which, according to Stich, expresses the doctrine that a computational theory of mind "ought to *restrict itself to postulating formal operations on mental states* [and] ought not to postulate processes which apply to mental states in virtue of their semantic properties" (Stich 1983, p. 162). This principle, according to Stich, "entails the rejection of the Strong RTM"; for, as he explains, "the theorist who couches his generalizations . . . in terms of the *content sentences* used to characterize mental states is merely postulating mental operations whose specification requires reference to semantic properties of these states" (ibid.)

As understood by Stich, methodological solipsism seems indeed incompatible with a "semantically based" theory like the strong RTM. But is methodological solipsism, as understood by Stich, a reasonable constraint on computational theories? Is it one to which the RTM is committed?

In the version made popular by Fodor (1980), methodological solipsism asserts that mental processess "have no access to the semantic properties of representations, including the property of being true, of having referents or, indeed, the property of being a representation of the environment" (Fodor 1980, p. 64). But what this means, exactly, and what methodological constraints this doctrine imposes on psychological theories, is far from obvious. The following considerations are in order.

1. In Putnam's (1975) original use of the term, methodological solipsism states the require-

ment that psychological theories should only admit psychological states in the *narrow sense*—where a psychological state in the narrow sense is one which need not satisfy such semantic properties as truth and reference in order to be truly ascribed. (Thus *believing that-p* is a psychological state in the narrow sense, whereas *knowing that-p* is a psychological state in the *wide* sense.) As distinct from epistemology, a (solipsistic) cognitive psychology *need* only admit narrow psychological states; knowledge, as distinct from "mere" (firmly held) belief, contributes nothing further to the etiology of behaviour.

2. It is now customary to distinguish between *referential (truth conditional)* semantics and *functional (conceptual, inferential) role* semantics. Whereas truth and reference find their natural home in a referential semantics, *meaning* and *content* may be understood either in terms of a referential semantics, or in terms of a functional-role semantics—in terms of the functional-conceptual-inferential role that items to which those properties are ascribed play in a system, given certain inputs. When the claim is made that methodological solipsism excludes reference to the semantic properties of the representations, including their content, the question ought to be asked: semantic properties in *which sense?*

It is tempting to answer: in the sense specified by a referential semantics, but not in the sense specified by a functional role semantics, for the latter does not presuppose anything about truth conditions, or about the way the world actually is as distinct from the way it is represented as being. It is far from clear, however, that this is the right answer. Fodor (forthcoming), for example, seems to favour a notion of mental content ("narrow content", he calls it) which presupposes a genuinely referential semantics—one, however, which only counts as distinct those objects,referents and truth conditions which are phenomenologically distinct and which can eventuate in difference of computational-behavioural states in a perceiving organism.

This is not the place to try to settle this question. Suffice it to say that the notion of content required by a semantically based computational psychology will have to be appropriately constrained so as to satisfy the intuition that cognitive behaviour is necessarily a function of the way an organism *represents* the world, but not necessarily a function of the way the world *actually* is. Call content thus constrained "narrow" content, without, however, committing ourselves to Fodor's specifically referential construal of this notion. Does then methodological solipsism exclude reference to the (narrow) content of mental representations, thus committing us to a purely syntactic theory of the mind?

Fodor's view is that adopting methodological solipsism is tantamount to adopting the *formality condition*, which is simply the condition that every cognitive operation must be formally representable. The point of adopting the condition is just to insure that any posited psychological operation is *computable*, thus *machine executable*, and thus causally and physically *realizable*; for a machine or physical system can only detect the formal features—the "shapes"—of the objects on which it operates. In no way does the formality condition prevent a theorist from ascribing such properties as meaning or content (*narrow* content, that is) to the constructs of his theory, if doing so enables him to recover theoretically useful generalizations about the objects in the domain of his theory; the formality condition only requires that the relevant semantic properties be *formally representable* and thus detectable by a machine.

In so far as the principle of methodological solipsism is understood merely as a consequence of the formality condition, there is no ground to Stich's supposition that it "entails the rejection" of a semantically based computational psychology of the strong RTM type. It is a mistake for Stich to construe the formality condition as requiring that "semantic properties of mental states play no role in the specification of psychological generalizations" (Stich 1983, p. 162), if meaning or (narrow) content count as semantic properties. The formality condition only requires that a computational theory tell us in virtue of what formal (and thus physically executable) operations certain generalizations statable over contents are satisfied. As I see it, the formality condition simply amounts to a *realizability constraint* on theories: a constraint demanded by a physicalistic ontology.

As already indicated, Stich of course does not deny that mental representations may *have*

content; he only holds that because of the formality condition generalizations stated over contents, even when available, will "do no work" in a computational theory, for extensionally equivalent generalizations statable over purely syntactic objects will always be available. But this will be true only on the assumption that a type-type reduction of contents to forms is possible, and there is no reason to suppose that such an assumption is true. The formality condition, at any rate, only requires that for any content-type there be at least one syntactic token and that distinct content-types have distinct syntactic tokens.

In conclusion, whether a cognitive theory should posit mental states with intentional contents would seem to depend on whether an individuation of such states in terms of their content helps us recover theoretically important generalizations about cognitive behaviour that we wouldn't otherwise recover; it does *not* depend on whether we accept or reject the formality condition and the related principle of methodological solipsism. *(Of course* we must accept the formality condition, if we wish the processes subsumed by our generalizations to be computationally and physically realizable!) There is thus no need to rule out, on formal a priori grounds, the Strong RTM nor, indeed, the possibility of a "science of the intentional."

ENDNOTE

* I am grateful to the Social Sciences and Humanities Research Council of Canada for a travel grant that made presentation of this paper possible.

REFERENCES

Fodor, Jerry, "Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology", The Behavioral and Brain Sciences Vol. 3 (1980).
Fodor, Jerry, *Representations* (Cambridge, Mass. 1981).
Putnam, H., "The Meaning of 'Meaning' ", in K. Gunderson (ed.), *Minnesota Studies in the Philosophy of Science* Vol. 7 (Minneapolis 1975).
Stich, S., *From Folk Psychology to Cognitive Science* (Cambridge, Mass. 1983).

* * *

FUNCTION, LAW AND THE EXPLANATION OF INTENTIONALITY

Eric Russert Kraemer
University of Nebraska, Lincoln

In this paper I would like to consider two recent attempts to account, in a physical way, for the intentionality of mental states. The efforts I shall discuss are those of Fred Dretske and Berent Enç. Both seem to me to agree in holding the intentionality is to be explicated as derivate from certain other properties of physical systems. In what follows I shall (1) state the problem, (2) set out the views of the two authors, (3) assess their claims concerning intentionality, and (4) make some final suggestions.

1. If mental states are wholly physical items, then how can many of them (most notably, the propositional attitudes) be "directed upon" objects or states-of-affairs which need not exist? Let us call this the problem of intentionality. It is the difficulty of answering this question which has prompted some followers of Brentano to urge that materialism in any interesting form can not be correct. But, we may ask, what is any serious theorist of the mind-body relation to make of Brentano's claim (Brentano, 1874) that mental phenomena are to be distinguished from physical phenomena precisely because the former and not the latter can be directed in a special way towards objects? There seem to be several sorts of answers that may be given.

One response to make is to deny that this sort of "directedness" is indeed a serious ontological feature. (See, for example, Quine, 1960). A second way of answering the problem would be to attempt to provide an *extensional* account of intentional concepts. (See, for example, R. J. Nelson, 1982). A third approach to the problem of intentionality would be to argue that intentionality is a primitive, unanalyzable property of certain mental states. I shall not comment on any of these three manoeuvers here. I shall concentrate intead on a fourth way to answer the problem of intentionality which I think deserves serious consideration. This concern examining what have been alleged to be the features (or 'criteria') of intentionality and attempting to explain how these can be present in a physical system given a certain conception of the mental. Let us call this the *criterial* approach. There are two ways in which this strategy might be successful. One might, first of all, look to the important concepts of the theory of mind under consideration. One might be able to demonstrate that some of these concepts themselves exhibit intentionality. The intentionality of the mental would then be explainable as derivative from that of certain concepts of the theory. There might, of course, be further worries in terms of accounting for the intentionality of these concepts, but if the theory of mind being discussed is non-circular, at least the problem of the intentionality of the mental would have been solved. A second way to implement the criterial approach would be to provide a set of explanations in terms of the elements of the theory of mind being proposed which would account for the presence of the intentional features of the mental system, even though the elements of the explaining theory lacked those features themselves. Let us now consider how the criterial strategy has been applied.

2. Fred Dretske's approach to mental states is by way of information thery. (Dretske, 1980) He holds that cognitive states should be construed as involving the transmission, receipt and processing of information, where information transfer depends upon nomic connections. Fred Dretske says: ". . . the intentionality associated with our cognitive states can be viewed as a manifestation of any underlying network of nomic regularities" (Dretske, p. 287). Dretske's argument for this claim consists in his urging that nomic regularities (in terms of which he

analyzes cognitive capacities) *as* nomic regularities (or laws of nature) possess "a significant degree of intentionality" (Dretske, p. 290). Enç, on the other hand, holds a view that he calls *teleological functionalism*. Enç identifies psychological states with the functional states of psychological mechanisms. Enç states

> ". . . The intensional language in which descriptions of psychological states are couched can be shown to be generated by the implicit function attrubutions made to, and by the functional explanations given for psychological phenomena, and when states of physical systems are referred to in the context of such functional explanations, it can also be shown that the language in which these states are described has the same degree of *intentsionality* as that of the language of psychology." (Enç, 1982, pp. 181–182)

(A similar claim is suggested, though not argued for, in William Lycan, 1981.) Let us now turn to assessing the claims of these two authors.

3. Let us now investigate to what extent we can ascribe intentionality to nomic and functional statements. Let us take as basic schemata the following:

(L) It is a law of nature that all F's are G's.
(F) A function of X is to bring about Y.

Specific instances of (L) and (F) do seem to satisfy a number of the criteria that have been proposed for intentionality.

Let us now begin with an important early criterion for intentionality, namely, non-substitutivity of co-referential expressions. (Chisholm, 1957) According to this criterion, if a sentence contains some term *t* such that one cannot substitute a co-referential term for *t* without changing the sentence's truth value then the sentence is intentional. Dretske urges, I think correctly, that nomic statements do satisfy this criterion:

> "To use a well-known example, drunks have liver problems and their problems may have their basis in a nomic relationship between excessive alcoholic intake and the condition of the liver. But we cannot infer from this fact that there is a nomic connection between sitting on a park bench B and liver trouble just because all, only only, drunks sit (have sat, are sitting, will sit) on park bench B." (Dretske, p. 286)

And Enç has claimed, persuasively, that the criterion also applies to function statements. He offers the following example:

> "Some fruits flies are known to be more fit in environments of high himidity. These fruits flies have a sensor which in their natural environment helps them to seek humid spots. The sensor functions by detecting the intensity of light in environment and helping the organism to steer itself to darker areas. Since in the natural environment of these flies there is a law governed correlation between dark aeas and humid areas, the organisms by and large end up in the more humid spots . . . A good case can be made for the claim that the function of the sensor is to enable the organism to reach humid spots, and not to enable the organisms to reach dark spots." (Enç, p. 168)

Dretske and Enç do not explicitly argue that other criteria of intentionality can be explained by appeal to features of statement of the form of (L) and (F). So to assess their claims it is worth considering which of the other proposed criteria of intentionality are satisfied by law and function statement. A second proposed criterion of intentionality is the criterion of non-extensional occurrence. (Chisholm, 1957) According to this criterion, if a sentence contains a dependent clause *C* such that one cannot substitute any clause of the same truth-value for *C* and preserve the truth-value of the sentence then the sentence is intentional. Clearly sentences of the form of (L) will satisfy this criterion, for certain obvious substitution instances of the clause, all F's and G's, will produce not only false, but necessarily false statements. To show that the non-extensional occurrence criterion also to sentences of the form of (F) one might suggest that (F) is logically equivalent to

(F*) X has the function of bringing it about that Y occurs.

If this equivalence is granted, then it is a simple matter to show that particular substitution instances for the clause 'Y occurs' will conform to the criterion.

A third criterion to consider is that of failure of existential generalization. (Chisholm, 1957) According to this criterion a sentence is intentional if its truth value implies neither the existence nor non-existence of the things it talks about. I assume it obvious that instances of (L) satisfy this criteria. For example, laws relating the melting points of golden mountains imply neither that there are nor that there are nor any such mountains.

The case for (F) seems a bit more problematic. One might say that sentences of this form use the expression 'Y' in such a way that neither (F) nor its contradiction implies that there is or is not anything to which 'Y' applies. For it seems that items may certainly have functions *without* its being the case that those functions are successfully carried out. The objection may be made, however, that ordinary examples fail to bear this out. Take, for instance,

The function of the pancreas is to secrete insulin.

Individual pancreases may fail to produce insulin; nevertheless, insulin still exists. To make the argument for the failure of existential generalization criterion's applying to function statements stronger, let us consider what I will call the case of the unfortunate mutant. Imagine two organism 0_1 and 0_2 (the only two of this type) each with a *spancreas* whose function (it is later revealed) is to produce *spinsulin*-a type substance no instance of which ever *exists-but* which can be inferred from chemical and biological theory. But suppose that 0_1 and 0_2's spancreases never matured sufficiently to produce spinsulin for poor 0_1 and 0_2 were caught in a passing lava flow before having a chance to reproduce. Now according to one way of conceiving the matter, an organ such as a spancreas could not have had the function of producing spinsulin if no spinsulin were ever produced. (This result seems required, for instance, by Larry Wright's account of functions, Wright, 1973) But I do not see why one should be forced to accept this conclusion. It may be that in order for us to be able to establish conclusively a functional relationship we would need to refer to some existing state. But to argue that this epistemic handicap shows that all goals of functional states must be instantiated a t least once seems to me to confuse epistemology with ontology.

I have now considered three criteria that have been proposed for intentionality and have suggested that they apply as well to nomic relations and function ascriptions. There are, however, certain features of intentionality that it is more problematic to attribute to nomic and function ascriptions. Let us consider two different examples. Let us begin with what Chisholm has called a *negative* criterion of intentionality. (Chisholm, 1966) Suppose we consider the belief that *p*. Let us represent this as 'Bp'. We can then express the following claims as characteristic of the logic of belief statement: (A) $Bp \nrightarrow \sim p$; (B) $Bp \nrightarrow \sim B \sim p$; (C) $\sim B \sim p \nrightarrow \sim B\, p$; (D) $\sim B \sim p \nrightarrow Bp$. This can be generalized as a criterion of intentional modality. How do nomic and functional statements fare according to this criterion? Although analogues of (A), (C) and (D) are applicable to such statements, there seems to be a problem with finding correct analogues for (B). That is, its being a law of nature that all F's are G's does seem to entail that it is not a law of nature that not all F's are G's; and, its being a function of X to bring about Y does seem to entail that it is not a function of X to bring about $\sim$ Y. Now one response here is to question whether (B) does hold for belief, for (B) requires the possibility of holding inconsistent beliefs, a possibility which is most controversial. (See, e.g., the paper by J. J. Ross in this volume.) Without going into that debate let me note two points. First, there are other mental attitudes, such as desire, which do seem to be strong candidates for satisfying this criterion. And secondly, I think it at least an *issue* whether inconsistent beliefs are possible, whereas there seems to be no such issue regarding function statements or laws of nature.

I would like to mention one further criterion of intentionality which I think is also problematic. Chisholm has suggested the following:

> . . . a sentence prefix M is intentional if, for every sentence q, the result of modifying q by M is logically contingent; and a sentence is intentional if it is the result of modifying a sentence by an intentional prefix or if it implies a sentence that is intentional. (Chisholm, 1966, p. 54)

Let us call this the *contingent prefix* criterion. Chisholm suggests that psychological states unlike all others can only be described using sentences which are intentional in this sense. Let us consider again the sentence schema (F*). The appropriate sentence prefix of this is "X functions to bring it about that." The problem that the contingent prefix criterion raises for the teleological functionalist is as follows. Consider any simle necessary arithmetic truth such as "2 + 2 = 4". If we prefix the sentence prefix of (F*) to such a sentence, the result, "X functions to bring it about that 2 + 2 = 4", does not just seem to be false as a matter of accident for substitution instances of X, but rather it seems to be *necessarily* false. Thus the contingent prefix criterion cannot be easily satisfied by teleological prefixes. Analogous remarks are to be made concerning nomic prefixes. But then how is this aspect of intentionality to be explained by those who use the criterial strategy?

5. I will now summarize what I have argued and make some final comments. Dretske and Enç are right in supposing that certain aspects of intentionality can be attributed to certain physical properties. Where they go wrong is in underestimating how *many* aspects of intentionality (as opposed to intensionality) so not readily apply to these notions. I have discussed only two such aspects, but there are many more to be considered. (See, for example, contributions by H. Morick and R. M. Chisholm in this volume.) The conclusion to draw is that the task of providing a physical explanation of intentionality is much farther from having a solution than Dretske and Enç confidently maintain.

My discussion here has dealt only with the first of the two ways I suggested for implementing the criterial strategy. However, it seems to me that if the materialist is to make significant progress on the problem of accounting for intentionality, then the second and more difficult procedure for carrying out the strategy needs to be explored.

One final aside. Some philosophers may take the attempt of the past thirty years to provide a mark of intentionality to have been shown to be an unproductive enterprise, with the literature being littered with only unsuccessful criteria. Although these various criteria have not been widely perceived to have succeeded in singling out all and only mental phenomena, nonetheles they do give us a useful benchmark by means of which we may gauge which aspects of intentionality can be explicated by the various physical theories of the mind.

REFERENCES

Brentano, F., *Psychology from an Empirical Standpoint* (Leipzig 1874), Book II, Ch. 1.
Chisholm, R. M., *Perceiving* (Ithaca 1957), Ch. 11.
Chisholm, R. M., "On Some Psychological Concepts and the 'Logic' of Intentionality", and "Rejoinder", in: H. N. Castañeda (ed.), *Intentionality, Minds and Perception* (Detroit 1966), 11–35, 46–52.
Dretske, F., "The Intentionality of Cognitive States", *Midwest Studies in Philosophy*, Vol. 5 (1980), 281–294.
Enç, B., "The Intentional States of Mechanical Devices", *Mind*, Vol. 94 (1982), 161–182.
Lycan, W., "Form, Function and Feel", *Journal of Philosophy*, Vol. 78 (1981), 24–50.
Nelson, R. J., *The Logic of Mind* (Dordrecht 1982).
Quine, W. V., *Word and Object* (Cambridge, Mass. 1960), section 45.
Wright, L., "Functions", *Philosophical Review*, Vol. 82 (1973), 139–168.

* * *

DIE INTENTIONALITÄT DES DENKENS

Wolfgang Künne
Universität Hamburg

Wann ist das, was einer denkt, dasselbe wie das, was ein anderer denkt, – oder wie das, was er selbst zu einer anderen Zeit denkt? Das ist die Frage, die ich – im Blick auf ein Denken, das mit indexikalischen Aussagesätzen kundgetan werden kann – erörtern möchte. Mit ‚denken' ist dabei gemeint: denken, *daß* . . ., und zwar episodisch im Sinne von ‚urteilen' und dispositional im Sinne von ‚glauben'.

Dort, wo die Analytische Philosophie der letzten Jahre auf unsere Fragestellung eingegangen ist, ist eine Kategorie zu neuen Ehren gekommen, die Russell und Whitehead eingeführt haben: die Kategorie der *singulären Propositionen*. Ich werde zunächst diese Kategorie erläutern. Sodann werde ich dafür plädieren, die singulären Propositionen, die durch indexikalisch kundgegebene Urteile und Meinungen repräsentiert werden, zu unterscheiden von den *Inhalten*, die solche Urteile und Meinungen haben. (Diese Distinktion ist von Husserl und Reinach inspiriert.) Abschließend werde ich versuchen, das kritische Potential dieser Unterscheidung anzudeuten.

Wenn ein Satz einen Indikator enthält, so ist das, was der Indikator in einem gegebenen Kontext denotiert, ein „Bestandteil" (constituent) der in diesem Kontext ausgedrückten Proposition. Nimmt also jemand mit dem Indikator in einer Äußerung von

(S1) Der da ist stupsnasig

auf Sokrates Bezug, dann ist Sokrates selber (in propria persona) ein Bestandteil der ausgedrückten Proposition. Die andere Komponente dieser Proposition ist die Eigenschaft, stupsnasig zu sein. So bestimmen Russell, Whitehead und neuerdings Donnellan, Kaplan und Perry die Struktur einer singulären Proposition (kurz: einer sP).[1]

Jeder sP entspricht genau ein geordnetes Paar, dessen erstes Glied ein Einzelding oder ein geordnetes n-tupel von Einzeldingen und dessen zweites Glied eine Eigenschaft oder eine n-stellige Relation ist. So können wir der sP, die mit (S1) bei Bezugnahme auf Sokrates ausgedrückt wird, das geordnete Paar (Sokrates; Stupsnasigkeit) zuordnen. Aber nicht jedem solchen Paar entspricht eine sP. So können wir z.B. der folgenden Sequenz keine sP zuordnen: ((Sokrates; Simmias); die QUADRAT-VON-Relation). Die Spezifikation einer sP soll nämlich eine mögliche Antwort auf die Frage sein, was einer glaubt oder urteilt. ‚A glaubt, daß dies (→ Sokrates) das Quadrat von diesem (→ Simmias) ist' kann aber keine Antwort auf die Frage sein, was einer glaubt. (In diesem Sinne hat Wittgenstein 1913 gegen Russells damalige Urteilstheorie eingewandt: Eine richtige Theorie des Urteils muß zeigen, daß es unmöglich ist, einen Unsinn zu urteilen [*TLP* 5.5422].) SP sind also etwas anderes als geordnete Paare des angegebenen Typs.[2] Was die *Darstellung* von sP durch geordnete Paare augenfällig machen soll, ist die Tatsache, daß die Identität einer solchen Proposition durch die von den Indikatoren denotierten Gegenstände *selber* bestimmt wird – und *nicht* durch die Art und Weise, wie diese Gegenstände dem Denkenden gegeben sind.

Wir können Urteilsakte und Meinungen nun unter dem Gesichtspunkt klassifizieren: *Welche von ihnen repräsentieren dieselbe singuläre Proposition?* Ein mit einem indexikalischen Satz S kundgegebenes Denken repräsentiert genau dann dieselbe sP wie ein mit dem indexikalischen Satz S' kundgegebenes Denken, wenn es jeweils genau eine Entität gibt, auch die mit den Indikatoren in S und mit ihren Gegenstücken in S' Bezug genommen wird, und wenn über diese Entitäten mit S und S' dasselbe gesagt wird.

Nun kann aber manchmal der ernsthaften assertorischen Äußerung eines indexikalischen Satzes *keine* sP zugeordnet werden. Macbeth halluziniert und schreit:

(S2) Das da ist ein blutiger Dolch.

In Ermangelung eines Gegenstandes, der in dieser Situation durch den Indikator denotiert wird, kann man nicht sagen, daß mit (S2) eine sP ausgedrückt wird. Macbeths Äußerung von (S2) ist aber gewiß ein guter Grund, ihm einen Urteilsakt zuzuschreiben. Nun kann man nicht urteilen oder meinen, ohne etwas zu urteilen oder zu meinen. *Was* ist das aber in diesem Falle? Es gibt etwas, das Macbeths kognitive Situation mit derjenigen gemeinsam hat, in der er sich befunden hätte, wenn ihm die indexikalische Bezugnahme *nicht* mißglückt wäre. Sein Denken hätte dann (ceteris paribus) denselben bedeutungsmäßigen Inhalt gehabt.

Wir können Urteilsakte und Meinungen nämlich auch unter dem Gesichtspunkt klassifizieren: *Welche von ihnen haben denselben bedeutungsmäßigen Inhalt?* Urteilsakte und Meinungen haben genau dann denselben bedeutungsmäßigen Inhalt, wenn sie durch Aussagesätze kundgegeben werden können, die denselben lexikalisch-grammatischen Sinn haben.

Die Zweideutigkeit des folgenden Satzes reflektiert die Verschiedenheit der Hinsichten, die uns bei der Bestimmung dessen, was einer denkt, leiten können:

(S3) Am Vorabend der Wahl dachte A, er werde die Wahl gewinnen,
und sein Gegenspieler B dachte *es* auch.

Was dachte B auch? In der *einen* Lesart wird die sP, die B's Denken repräsentieren, mit derjenigen identifiziert, die A's Denken repräsentiert: Beiden denken von A, er werde gewinnen. In der *anderen* Lesart wird der bedeutungsmäßige Inhalt, den B's Denken hat, mit demjenigen identifiziert, den A's Denken hat: Jeder denkt, er selbst werde gewinnen; jeder könnte seine Meinung mit dem Satz ‚Ich werde die Wahl gewinnen' (oder mit einem synonymen Satz) kundtun.

Manchmal benötigen wir einen reicheren Inhaltsbegriff als den des bedeutungsmäßigen Inhalts. Wenn jemand auf Hinz zeigt und, ohne den Finger zu bewegen, sagt:

(S4) Dies ist derselbe Mann wie dies,

so trifft er eine nicht sonderlich erhellende Feststellung. Aber wenn jemand mit Hinz telefoniert, dabei aus dem Fenster sieht und Hinz in der Telefonzelle vor dem Haus erblickt, so kann er mit (S4) sehr wohl eine erkenntniserweiternde Aussage machen. In beiden Fällen repräsentiert das Denken des Sprechers dieselbe sP, der man das folgende geordnete Paar zuordnen kann: ((Hinz; Hinz); Identität). Aber in beiden Fällen hat das Denken *auch* denselben bedeutungsmäßigen Inhalt. Die Nicht-Trivialität dessen, was im zweiten Fall gedacht wird, kann also nicht unter Rekurs auf den *bedeutungsmäßigen* Inhalt des Urteilsaktes verständlich gemacht werden. Zum *vollen* Inhalt gehört hier eben auch noch, wie dem Denker das Objekt seiner indexikalischen Bezugnahme jeweils perzeptiv gegeben ist.[3]

Zwei Beispiele mögen abschließend das kritische Potential der skizzierten Distinktion sichtbar machen.

(a) Kann man Unmögliches glauben?[4]

‚Es ist trivial falsch, daß p' möge genau dann gelten, wenn niemand ‚p' verstehen kann, ohne eo ipso zu wissen, daß es falsch ist, daß p. Wenn ‚p' diese Bedingung erfüllt, dann kann in der Tat niemand glauben oder urteilen, daß p. Der lexikalisch-grammatische Sinn von ‚p' kann dann nicht der bedeutungsmäßige Inhalt eines Glaubens oder Urteilsaktes sein. Sätze von dieser Art sind z.B. ‚2 ist eine Primzahl, und 2 ist keine Primzahl', ‚Alle Junggesellen sind weiblich', und

(S5) Dies ist von sich selbst verschieden.

Aber ein Satz ‚p' kann auch dann etwas ausdrücken, was unmöglich der Fall sein kann, wenn es nicht *trivial* falsch ist, daß p. Und dann kann jemand sehr wohl glauben oder urteilen, daß p.

Wenn nun in einer Äußerung von (S4) beide Indikatorenvorkommnisse Hinz denotieren, dann drückt (S4) etwas aus, was notwendigerweise der Fall ist. Und wenn in einer Äußerung von

(S6) Dies ist ein anderer Mann als dies

wieder beide ‚dies' Hinz denotieren, dann drückt (S6) etwas aus, was *unmöglich* der Fall sein kann. Ist die indexikalische Bezugnahme auf Hinz dabei aber einmal mit einer auditiven und einmal mit einer visuellen Hinz-Wahrnehmung verknüpft, so kann man mit (S6) sehr wohl einen Urteilsakt kundtun. Bei Bezugnahme auf Hinz drücken nun (S6) und (S5) dieselbe sP aus; wir können ihr das geordnete Paar ((Hinz; Hinz); Verschiedenheit) zuordnen. Aber der Sinn von (S6) kann sein, was der Sinn von (S5) nicht sein kann: der bedeutungsmäßige Inhalt eines Glaubens oder Urteilsaktes.

(b) Gibt es indexikalisch ausgedrückte Gedanken, die kontingent und dennoch a priori sind? Gibt es indexikalisch ausgedrückte Gedanken die notwendig, aber gleichwohl a posteriori sind? (Ersetzt man ‚indexikalisch' durch ‚unter Verwendung von Eigennamen', so lautet Kripkes Anwort auf beide Fragen bekanntlich: Ja.)

„Es ist kontingent, daß ich jetzt hier bin; aber (so könnte jemand argumentieren) daß ich jetzt hier bin, weiß ich a priori." – Diese Erwägung ist verworren. Jeder, der den deutschen Satz

(S7) Ich bin jetzt hier

versteht, weiß, daß mit ihm (in seiner gestenunabhängigen Standardverwendung[5]) stets etwas Wahres gesagt wird. Ich verstehe (S7). Also kenne ich *a priori* den Wahrheitswert des *Inhalts*, den das jetzt von mir mit (S7) kundgegebene Urteil hat. Aber es ist in der Tat nicht notwendig, daß ich jetzt hier bin; schließlich hätte ich ja auch in meinem Geburtsort bleiben können: Die *sP*, die mein Urteil repräsentiert, ist *kontingent*. Die epistemische Eigenschaft, a priori von mir als wahr erkannt zu werden, und die modale Eigenschaft, kontingent zu sein, haben gar nicht denselben Träger.

„Daß dies" (der Sprecher hört Hinz) „derselbe Mann ist wie dies" (der Sprecher sieht Hinz), „ist notwendig; aber ich weiß es a posteriori." – Auch diese Erwägung ist verworren. In der geschilderten Situation ist die ausgedrückte *sP* Träger der modalen Eigenschaft *Notwendigkeit*, während der Träger der epistemischen Eigenschaft, vom Sprecher *a posteriori* als wahr erkannt zu werden, etwas anderes ist, nämlich der *Inhalt* seines Denkens.

AMERKUNGEN

[1] Vgl. W. Künne, „Indexikalität, Sinn und propositionaler Gehalt", *Grazer Philosophische Studien*, Bd. 18 (1982), S. 58f.

[2] Dies u.a. gegen S. Schiffer, „The Basis of Reference", *Erkenntnis*, Bd. 13 (1978), S. 171.

[3] Vgl. A. Woodfield, „On Specifying the Content of Thought", in ders. (ed), *Thought and Object* (Oxford 1982), S. 286.
Die in diesem Absatz angestellte Überlegung zeigt übrigens, daß Kaplans ganz anders motivierte Unterscheidung zwischen „content" und „character" für unsere Zwecke unzureichend ist.

[4] Nein, – antwortet R. Barcan Marcus, „Rationality and Believing the Impossible", *Journal of Philosophy*, Bd. 80 (1983). Das Folgende ist z.T. kritisch auf diesen Aufsatz bezogen.

[5] Zeigt jemand auf eine Landkarte und äußert dabei (S7), so kann er natürlich etwas Falsches behaupten.

* * *

INTENTIONALITÄT BEI KANT*

Hjördis Nerheim
Universität Oslo

1. Nach Kant verhalten wir uns zur Welt, in der wir leben, denken und handeln, in der Form von Urteilen. Daß wir überhaupt *urteilen*, ist für Kant gleichbedeutend mit der Tatsache, daß wir überhaupt *Bewußtsein* haben. Dies deshalb, weil das Urteil die Form der Wirklichkeit im Bewußtsein ist, in der der Mensch sich als solcher verwirklicht und um sich und seine Welt in der transzendentalen Apperzeption weiß. Mit anderen Worten: Kant behauptet, daß Selbstbewußtsein Urteilsfähigkeit voraussetzt.

Das Vermögen selbst zu urteilen oder die Urteilskraft ist nun nach Kant „ein Verhältnis, das objektiv gültig ist" (*Kritik der reinen Vernunft* (*K.d.r.V.*,) B 124). Urteilsfähigkeit ist also das Vermögen zur Objektivität bzw. Intersubjektivität, eher gar kein Vermögen, sondern der Gebrauch, den wir von allen unseren Vermögen in handelnder Absicht machen können. Sie ist die Tatsache, daß man Urteile bilden kann.[1] Als solche ist sie ein Können. Wer etwas kann, ist *frei* es zu tun. Somit gewinnt das Urteilen den Aspekt der Urteils*handlung*. Die Freiheit der Urteilskraft ist also letztlich Bedingung dafür, daß *überhaupt* gedacht oder geurteilt werden kann. Hierfür ist im Folgenden der Nachweis zu führen.[2]

2. Im Hinblick auf unser Thema steht zunächst fest, daß Kant erst in der *Kritik der Urteilskraft* (*K.d.U.*) klar zum Ausdruck bringt, von welcher Art die Urteilshandlung ist, die er bereits für die Rationalität des Erkennens und dann auch für die des Handelns in Anspruch genommen hatte: Kognitives Urteilsvermögen, so konstatiert Kant jetzt, ist eine „auf Erkenntnis gerichtete Absicht" (*K.d.U.*, XXXVIII); es ist mit anderen, wiederum Kantischen Worten, „unsere absichtliche Tätigkeit", womit die „Erkenntniskräfte ins Spiel gesetzt werden" (*K.d.U.*, 28). Oder anders ausgedrückt: Theoretisches Erkennen und erst recht dann sittliches Handeln ist *Intentionalität*. Menschliche Rationalität ist als theoretische oder praktische Urteilsfähigkeit recht eigentlich Intentionalität.

Von daher besteht die Aufgabe, die stillschweigenden Implikationen zu entfalten. Dies bedeutet als erstes sich folgendes klarzumachen: Ist kognitive Urteilshandlung recht eigentlich Intentionalität, so muß sie damit auch etwas sein, das entweder Erfolg oder Mißerfolg hat, das entweder glückt oder mißglückt. Infolgedessen kann sich auch als Erfüllung einer Intention immer nur etwas einstellen, was im Prinzip als Mißerfolg derselben ausbleiben könnte. Diese Feststellung impliziert, daß sofern wir nur immer etwas intendieren, dieses intendierte Etwas sich eben entweder als Verwirklichung dieser Intention einstellen oder als Mangel an Verwirklichung derselben ausbleiben muß. Dennoch ist Urteilsvermögen als Intentionales nicht in ein und demselben Sinne entweder erfolgreich oder erfolglos. Denn so etwas wie eine Intention intendiert niemals etwa, erfolglos zu sein, sondern ausschließlich die *geglückte* Erreichung ihrer Absicht. Daß die Erfüllung ausbleiben kann, ist *unintendiertes* Handeln, d.h. Täuschung, und dies ist *Mangel* an Erkenntnis und bereits ein Fall von Unfreiheit. Wir dürfen demnach sagen, daß Erkennen als Intention *ursprünglich* und unmittelbar immer nur Erfolgsintention wäre und niemals etwa Falschheitsintention. Folglich muß dann das, was wir als Wahrheit bzw. Falschheit von Erkennen kennen, streng nominalistisch als Glücken oder Mißglücken der immer nur auf *Wirklichkeit* gehenden Erkenntnisintention verstanden werden. Wirklichkeit selber ist nichts als *Selbstverwirklichung* des Subjekts und muß ebenso wie der Terminus ‚Bewußtsein' von dem Urteilen als freie Handlung her verstanden werden. Was wir darunter zu verstehen haben, wird erst später ausgeführt werden können. Doch steht schon jetzt außer Zweifel, daß Glücken und Mißglücken von Intentionalität notwendig mit zu ihrem Begriff dazugehört.

In dieser Perspektive hat Erkennen Gelingensstruktur. Und eben dadurch ist Erkenntnis als erfolgreiche Intention gerade nicht ausschließllich in jemandes Macht. Vielmehr ist sie, mit Kant zu sprechen, eine *Gunst*. Das Spezifisch-Intentionale einer kategorialen Urteilshandlung ist zwar notwendig, aber deshalb nicht auch schon notwendig erfolgreich, was Kant mit der These zur Sprache bringt, daß „der Verstand damit [mit den kategorialen Begriffen bez. Urteilsformen, H.N.] unabsichtlich nach seiner Natur notwendig verfährt" (*K.d.U.*, XXXIX). Dies besagt jedoch keineswegs, es finde dabei überhaupt keine Absicht statt. Vielmehr ist der eigentliche Sinn jener Behauptung der Unabsichtlichkeit des Verstandes bezüglich seiner Kategorien mit jener Einsicht in die Intentionalität von Erkennen und Handeln nicht nur vereinbar; es gehören diese Absichtlichkeit und jene Behauptung der Unabsichtlichkeit sogar notwendig zusammen. „Die gedachte Übereinstimmung [. . .], muß nach aller unserer Einsicht als *zufällig* [von mir hervorgehoben] beurteilt werden, gleichwohl aber doch für unser Verstandesbedürfnis als unentbehrlich, mithin als Zweckmäßigkeit, wodurch die Natur mit unserer, aber nur auf Erkenntnis gerichteten, Absicht übereinstimmt" (*K.d.U.*, XXXVII). Kant formuliert deren Zusammengehörigkeit als *Zweckmäßigkeit ohne Zweck*.

Da nun die Zweckmäßigkeit *ohne* Vorstellung eines Zwecks sein soll, vermag sie nur unmittelbar empfunden zu werden. Und dies überzeugt um so mehr, als nach Kant die Erreichung jeder Absicht mit dem Gefühl der Lust verbunden ist (*K.d.U.*, XXXIX). Und eben darum ist mit jedem Erfolg die Empfindung eines Glücks verbunden. Die Lust gehört notwendig zur Erkenntnis, aber nicht so, daß dabei etwas, aus welchen Gründen immer, gelänge, sondern vielmehr so, daß etwas sich darüber vollendet hat. Im kognitiven Bereich von Erkennen kann man die Berechtigung dazu vor allem aus dem Umstand beziehen, daß etwas glückt oder gelingt, was wir nicht *erzwingen* können. Die Lust liegt dabei nicht bei der Absicht oder Nichtabsicht, sondern bei der „Erreichung", d.h. dabei, daß etwas Neues zum Vorschein gekommen ist, was zuvor nicht da war. Und wenn kein Argument mehr sticht, diese *Lust* bleibt der stärkste Einwand gegen die Skepsis. Schließlich vermittelt uns jenes Gefühl der Lust – dies Kants Argument – den Grund „der allgemeinen Mitteilbarkeit unserer Erkenntnis, welche in jeder Logik und jedem Prinzip der Erkenntnis, das nicht skeptisch ist, vorausgesetzt werden muß" (*K.d.U.*, 66).

Ich will die Folgerungen, die aus dieser These zu ziehen sind, an dieser Stelle nicht weiter ausführen. Nur diesen Punkt möchte ich herausheben. Nur das Bemerkenswerte: Unhintergehbare Voraussetzung jeder Erkenntnis überhaupt ist bei Kant die Lebensform, die im Vermögen der Lust und Unlust verankert ist.

3. Daß die Lust überhaupt ausbleiben kann, setzt aber eine Selbstbestimmbarkeit des Subjekts voraus, die Kant als *Interesse* bezeichnet. Das Subjekt könnte schon in der „Wahl" seiner Urteilshandlung durch irgendwelche empirischen oder dogmatischen Einflüsse, z.B. durch die „Neigung" zu bestimmten Theorien bedingt sein. Das würde dann heißen, die Urteilshandlung folgte Gründen aus den Bedingtheiten des Subjekts. Urteilen als *freies* Handeln wäre nicht möglich.

Die Konsequenz, die sich daraus ergibt, ist bemerkenswert: Sie besagt ganz unmißverständlich: Unter diesem Gesichtspunkt könnte es nicht *selbstbewußt*, d.h. *objektiv* urteilen und damit im Sinne Kants gar nicht urteilen können. Meine Freiheit wäre schöner Schein. Nun mag ich gewiß von außen als Gegenstand des inneren oder äußeren Sinnes betrachtet – im Verhältnis zur Intentionalität des Subjekts also – immer in dieser Weise bedingt sein und damit gar nicht in der Lage sein, selbstbewußt Urteile zu fällen.

Das ist die Perspektive, die Kants Analyse des Urteils als freies Handelns eröffnet. Und eben darin ist die Freiheit des Menschen im Verhältnis zur *Intentionalität* des Subjekts die reine Kontingenz und Faktizität. Andererseits setzen wir uns aber im moralischen Bewußtsein als freie, d.h. zur Objektivität fähige Urteilsvermögen voraus. Und eben dadurch hängt die Wirklichkeit des Menschen auch noch von Bedingungen ab, deren Erfüllung kognitives Urteilsvermögen bei aller Macht seiner Intentionalität doch prinzipiell nicht mächtig ist.

Die Gründe für die hier nur angedeutete Aporetik sind in unserem Zusammenhang nicht ausführlich zu diskutieren. Sie liegen auch nicht so sehr auf dem theoretischen Felde. Sie lie-

gen eher und mächtiger im Phänomen der neuzeitlichen Weltentfremdung. Vielmehr soll die Frage zur Diskussion gestellt sein, was es eigentlich bedeutet, daß das Subjekt von sich selber her *voraussetzungslos* mit dem Urteilen anfangen können soll, also *objektiv* und vorurteilsfrei urteilen kann.

4. Die eigentliche Schwierigkeit liegt nun darin – und das bleibt ein unlösbares Dilemma – wenn man menschliche Rationalität offen als Intentionalität *allein* deklariert – zu zeigen, wie es möglich ist, daß das Bewußtsein mit sich selber und seinesgleichen seine Erfahrungen machen kann. Das Dilemma besteht darin, daß keine Intention als ihre Erfüllung etwa sich *selber* oder ihresgleichen intendieren könne. Denn so etwas wie eine Intention hat man nicht dadurch, daß man diese Intention selbst etwa allererst intendiert, sondern ausschließlich dadurch, daß man *irgendetwas*, und das heißt, irgendetwas *anderes* als diese Intention selbst intendiert, wobei die kognitive Urteilsintention eben das Objekt-Konstituierende selber ist. Das Bewußtsein setzt diesen Unterschied zwischen Subjekt und Objekt – Begehren und Begehrtsein – indem es weiß, daß es selbst als intentionales Urteilen das ist, was den Gegenstand konstituiert, d.h. denkt. Es ist *Gegenstands*bewußtsein, indem es den Gegenstand durch seine Urteilshandlung entstehen läßt. Dabei bleibt aber für eine Theorie dieser Intentionalität die Wirklichkeit des Bewußtseins, der Begriff des Bewußtseins vom jeweils Anderen seiner selbst. Sie ist zwar real, aber doch nur als die Realität des Bewußtseins.

Die Frage, die in diesem Zusammenhang gestellt werden muß, läßt sich demnach folgendermaßen formulieren: Wie machen wir den Schritt von der bloßen Erkenntnis der Gegenstände zum – nennen wir es einmal so – Bewußtsein der erfahrenden Instanz selber? Dies ist der Frage vergleichbar (eine Frage, die hier nicht geprüft werden kann): Unter welchen Bedingungen ist eine Person im Stande, sich auf sich *selbst* zu beziehen? Die Lösung dieses Problems soll nach Kant in der *K.d.U.* zu finden sein, nämlich wir vollziehen diesen Schritt durch das Gefühl der Lust und das darauf gerichtete Vermögen der *reflektierenden* Urteilskraft.

Offensichtlich um diesem Dilemma zu entgehen, unterscheidet Kant jetzt zwei Funktionen unserer Urteilskraft oder vielmehr zwei Arten des Urteilens: das bestimmende und das reflektierende, die wir nach heutigem Verständnis als exemplifizieren und identifizieren bezeichnen dürfen: „Ist das Allgemeine (die Regel, das Prinzip, das Gesetz [der Begriff]) gegeben, so ist die Urteilskraft, welche das Besondere darunter subsumiert [. . .], *bestimmend*“ (*K.d.U.*, XXVI). In der neuen Terminologie kann die Einsicht in die Intentionalität von Erkennen und Handeln folgendermaßen ausgedrückt werden: „Intendieren“, bzw. „kategorisieren“ heißt soviel wie: nach geeigneten Gegenständen gesucht und *vielleicht* fündig geworden sein. Der Urteilskraft wird dabei keine *Freiheit* eingeräumt, weil ihr der Begriff, demgemäß sie eine einzelne Handlung unter die Regel des im Begriff ausgesprochenen Gesetzes bringt, *vorgegeben* ist. Sie findet nur das, was die Regeln allgemeinheitsfähiger Suche zugänglich machen. Ist dies geschehen, dann ist für die Urteilskraft, die bestimmend verfährt alle Problematik erledigt. „Ist aber“, fährt Kant fort, „nur das Besondere gegeben, wozu sie [sc. die Urteilskraft] das Allgemeine finden soll, so ist die Urteilskraft bloß reflektierend“ (ibid.). Die reflektierende Urteilskraft ist das Prinzip der Erreichung eines Handelns, deren Verwirklichung nicht von einem verfügbaren Allgemeinbegriff erwirkt ist, sondern welche sowohl in bezug auf den einzelnen Menschen wie in bezug auf die Gemeinschaft von Menschen durch Handlungen erst hervorzubringen ist. Denn es kann kein vorweg gesetzter Begriff ausgemacht werden, der einer freien Urteilskraft einen allgemeinen Bestimmungsgrund geben könnte. Gäbe es einen solchen vorgängig im Bewußtsein liegenden Begriff, dann wäre er den Grund, auf welchen das Urteilen nach einer Regel folgen würde, d.h. das Urteilen als *freie* Handlung wäre nicht möglich. Daraus ergibt sich unmittelbar die Konsequenz, daß es keine wie auch immer geartete „Theorie“ gibt, die die Freiheit des Urteilens als Handlung bestimmen könnte. Und wir werden unsere These am besten dadurch erhärten, daß wir nun die Frage stellen, von welcher Art denn die Freiheit ist, die letztendlich Bedingung dafür ist, daß ein jeder für sich Erkenntnis und Selbstbewußtsein haben kann. Kant nennt sie „das freie Spiel“ der Erkenntnis- bzw. Urteilsvermögen, was er mit der These zur Sprache bringt, daß die Erkenntnisvermögen zu

einem „Erkenntnis überhaupt“ zu denken sind. Mit der Einführung dieses Begriffs ist gerade das gemeint, daß die Erkenntnisvermögen nicht als *bestimmte* Urteilshandlungen mit eindeutig bestimmten Intentionen, sondern *frei* als die Betätigung von Vermögen *als* Vermögen ins Spiel zu bringen sind. Ohne dies freie Spiel kann *überhaupt* nicht gedacht oder geurteilt werden. Statt selber nur absichtlich intendiert zu werden, ist das freie Spiel vielmehr gar nichts anderes als dieses absichtliche Intendieren selbst, das seinerseits wesentlich niemals ausbleiben kann, was vielmehr immer schon auftreten muß, damit sich überhaupt etwas, Erkenntnis, als *Lust* zu diesem Intendieren einstellen kann. Denken *besteht* in diesem freien Spiel, weil es für jede Erkenntnis mittels menschlicher, d.h. diskursiver Begriffe vorauszusetzen ist. Kant formuliert dies wie folgt:

> Der Gemütszustand in dem freien Spiele der Einbildungskraft und des Verstandes (sofern sie untereinander, wie es zu einem *Erkenntnis überhaupt* erforderlich ist, zusammenstimmen) – muß ebensowohl für jedermann gelten und folglich allgemein mitteilbar sein, als es eine jede bestimmte Erkenntnis ist, die doch immer auf jenem Verhältnis als subjektiver Bedingung beruht. (*K.d.U.*, 29)

Diese Freiheit im Spiel definiert sich somit *vor* dem Übergang in einer *bestimmten* Urteilsform. Auf das Urteilen als Handeln in Freiheit übertragen, heißt das, daß die Freiheit im Spiel unserer Erkenntnisvermögen als die reine Möglichkeit freien Handelns, besser: objektiven Urteilens immer schon, vorausgesetzt ist. Das schließt dann auch die Erklärbarkeit von Freiheit aus. Erklären kann man nur, was sich auf anderes nach kategorialen Begriffen zurückführen läßt. Die Freiheit im Spiel ist also auch nicht vom Willen intendierbar, weil sie Voraussetzung für alles Wollen ist, dem an einer Handlung gelegen ist. Die Urteilsvermögen befinden sich hier noch in einem freien Spiel mit ihren *Möglichkeiten*, und nur in diesem Stadium schließt sich dem Bewußtsein alle Wirklichkeit und es selbst erst als ein gegenständlich greifbarer Inhalt auf. Das fordert aber eine „Umkehrung gewohnten Denkens“, um sich *indirekt* der unhintergehbaren Voraussetzung *jeder Erkenntnis überhaupt* zu vergewissern, die unter allen intendierten Inhalten als solche nicht vorkommt, wenn sie sie auch alle trägt.

Dieser *Grund*, unter dem dem Menschen somit seine Freiheit erscheint, ist nach Kant das *Schöne*. Wir nennen also Gegenstände unserer Welt schön, aber wir beziehen uns offenbar *indirekt* auf die Voraussetzung jeder Erkenntnis überhaupt. Das mag paradox klingen, ist aber gar nicht rätselhaft, sobald man die Kontingenz und Faktizität der Welt, wie sie Kant im Verhältnis zur Intentionalität des Subjekts auslegt, berücksichtigt. Konkret gesprochen bedeutet das: Die Welt war schon schön, bevor unsere Aussagen über sie wahr sein konnten.

ANMERKUNGEN

* Der Norwegische Forschungsrat für Allgemeine Wissenschaften, NAVF, Oslo, ermöglichte den folgen Beitrag durch ein Reisestipendium.

[1] Man könnte das so verstehen, daß das Denken – das nach Kant immer ein Urteilen ist – auf die Urteilskraft – die Kompetenz aller möglichen Urteilsbildung – als auf die „transzendentalen Grammatik“ oder „Tiefengrammatik“ unseres Urteilens „höre“ und damit die bloße Subjektivität des Bewußtseins zum Überindividuellen hin überschreite.

[2] Auf eine detaillierte Darstellung von Kants relevanten Argumenten muß hier freilich verzichtet werden. Ich habe eine solche Darstellung versucht in meiner Habilitation „Das Konstitutionsproblem in Kants Kritik der ästhetischen Urteilskraft“. Oslo, 1985.

* * *

2. Geist und Gehirn

2. Mind and Brain

THE COSTS OF MECHANISTIC EMBODIMENT

Donald M. MacKay
University of Keele

As a working hypothesis, the idea that our conscious experience (and our mental activity in particular) is *embodied in* our brain activity offers attractions as an alternative both to classical interactionism on the one hand and to reductionist materialism on the other. I have elsewhere (1–5)* explicated the notion of embodiment by using the familiar example of the embodiment of an equation in a computer.

(a) There is a clear sense in which a computer programmed to solve Poisson's equation (say) has its behaviour *determined* by Poisson's equation and its given boundary conditions. Were this not the case, we could not take the output of the machine as a solution of that equation.

(b) There is an equally clear sense in which the computer has its behaviour *determined* by the laws of physics as applied to each of its physical components, taking due account of the interactions between these components, and of the constraints imposed on the whole physical system by its physical environment. Were this not the case, the designer would warn us that the machine could not be relied upon as a computer. His claim that its behaviour as a symbolic system is determined by Poisson's equation *depends upon* his assumption that its behaviour as a physical system is determined by the laws of physics. The two are not rival, mutually exclusive claims, but logically *complementary*. [6]

This example serves well enough to illustrate the kind of interdependence (rather than interaction) that I would denote by speaking of conscious cognitive agency as 'embodied' in, and determinative of, the workings of the agent's brain-and-body. It clearly provides a counterexample to a form of argument commonly employed against mechanistic theories of brain function. In particular, it throws doubt on the presupposition of many 'incompatibilists' that complete physical determination of human behaviour *must* be denied if we want to claim that that behaviour is (at least sometimes, to some extent) determined by conscious thinking, valuing and choosing on the part of the individual concerned. With respect to the hypothesis that conscious mental activity is embodied in brain activity in the relevant sense, this presupposition is clearly false. The example also illustrates well the need to distinguish between two kinds or levels of 'causal' connection—between *physical* causes on the one hand, and what systems engineers today term *informational* or *systemic* causes on the other. Physical explanations account for changes in a system in terms of the flow and exchange of energy and force; informational explanations do so in terms of the flow and exchange of information and control. The first trace the dependence of *force* on *force*; the second the dependence of *form* on *form*. Each legitimately uses the explanatory term 'because'; but much confusion arises if we fail to follow Aristotle's lead in distinguishing between the quite different senses of 'cause' involved. In particular, to speak of 'downward causation' between the informational and physical levels, as is done by Sperry[7] for example, can give rise to much misunderstanding, with its suggestion of some kind of convergence of forces from 'above' and 'below' upon the physical components, which Sperry describes as 'pushed and hauled about' by the 'causes' operating at the higher level.[8] It would seem oddly unhelpful to a computer scientist to speak of the transistors in his machine as 'pushed and hauled about' by the equation it is solving! The distinction needed here seems closely related to Chisholm's well-known contrast between 'immanent' and 'transeunt' causality,[9] though the implications I would draw are not in favour of Chisholm's incompatibilist position; but in case I mistake his meaning, I will here use 'systemic control' to refer to the determination of behaviour by the higher-level process embodied in a physical structure, and 'physical causation' to refer to the determination of the behaviour

of the physical components, both individually and collectively, by the physical forces acting upon them.

But although I have been arguing on the foregoing lines for the past 35 years, and believe the computer analogy to be valid in these specific respects, I fear that comparisons between brains and computers in general, and the growth of the discipline called 'Artificial Intelligence' (AI) in particular, have given rise to a crop of deep misunderstandings as to the philosophical implications of current efforts to explain human behaviour mechanistically. Among the 'costs' of mechanistic embodiment, it is widely supposed, would be the reduction of all human behaviour to rule-following, the downgrading of concepts of 'tacit knowledge' or the 'ineffable' to mere temporary tokens of the incompleteness of our knowledge of the human 'brain program', the denial of freedom of choice as a (perhaps necessary) illusion based on ignorance, and the rejection of all religious assessments of our spiritual significance and destiny as incompatible with scientific mechanism. Whether each of these would be regarded as a 'cost' or a 'benefit' may be a matter of individual judgment; but I believe all of them to be unwarranted by the hypothesis of mechanistic embodiment. The bulk of this paper will therefore be devoted to disposing of some mistaken inferences, as I see it, from the neuroscientific comparison between brains and computing machinery.

There are several respects in which my example of the embodiment of an equation in a digital computer is potentially misleading. First, in that case we start with our explicitly t formalized equation, and our task is to construct an explicit embodiment of the equation in a general-purpose discrete-state machine, each of whose states serves as a distinct symbol in the relevant mathematical calculus. In the case of conscious cognitive agency, however, we start with our immediate unquantified experience of what it is like to see, feel, act, think, doubt, suffer, love, hate . . ., and it is far from obvious that any explicit program of transitions between discrete symbolic states could even approximate to a representation of that experience. The assumption that our brain is a mechanism does nothing to foreclose this question, since as we shall see below there are other types of mechanism in which transitions between states are not determined by consulting explicit rules.

Secondly, there is a serious disanalogy between embodied equations and embodied persons with regard to the problem of *identity*. To ask whether the equation embodied in a particular computer is "the original equation or merely a replica" would be to raise a silly question. Poisson's equation is one and the same, no matter how many embodiments of it may be knocking around in people's computers. But in the case of you and me the question of personal identity is far from silly, however difficult we may find it to express in sharp and satisfactory terms.

Thirdly, whenever the appropriate program is run in a suitable machine, we can guarantee that its behaviour is determined by Poisson's equation; but it is far from obvious that whenever an AI program is run in a suitable machine so as to generate behaviour we can recognize as intelligent, we are entitled to conclude that we are in the presence of a conscious individual.[2] To repeat an argument I have used elsewhere,[10] suppose we visit the theatre, where Joe Smith is playing Hamlet in the famous soliloquy. How many people—conscious centres of awareness—are on the stage? Obviously, if Joe does his job, there is a sense in which what we can see is 'Hamlet's behaviour'; but this does not mean that Hamlet is present as a conscious centre of awareness in addition to Joe. Generating Hamlet-behaviour by consulting a book of rules does not conjure Hamlet into present existence as a conscious person. Even if Joe deserts Shakespeare's text, quits the stage and 'adopts the character' of Hamlet in free dialogue with us, it would be Pickwickian to claim on these grounds that we had met Hamlet, as distinct from an impersonation of Hamlet. It would be totally irrational to argue that the ontological situation might change if only Joe were given enough time to perfect his skill in generating Hamlet-behaviour. Merely organizing behaviour according to rules, whether explicit or implicit, is not a sufficient way to increase the number of centres of conscious awareness in the world.

If it would be irrational to take the Hamlet-like behaviour of Joe Smith as proof of the presence of Hamlet, it is not obviously less irrational (*pace* Turing[11]) to take the person-like behaviour of an AI-programmed computer as proof of the presence of a conscious person. That

such a computer can have a mind or personality, in the sense in which a character represented on the stage has a mind and personality, need not be disputed;[12] but that a computer programmed by rule to act the part of a person has become *ipso facto* the embodiment of an additional member of the cognitive community is a quite baseless notion, which no amount of refinement in the acting skills of artificial agents can render plausible. Impersonation is not the same as embodiment.

At this point the reader may be tempted to suppose that our argument tells equally against the hypothesis that the brain is explicable in mechanistic terms; but this would be a mistake. The common notion that all machines are formally equivalent to some explicit program is false, and the truth is that only a restricted class of machines are rule-followers in the foregoing sense. 'Intelligent' machinery is characterized by the presence of branch-points in its course at which one out of several possibilities must be selected. In digital computers of the sort normally used for AI, the outcome at each branch point is determined by consulting a rule, which may involve extremely complex computations, but is explicitly specified by the program. In other types of mechanism, however, a quite different principle is employed: the outcome is determined, at least in part, by what amounts to the *making of a physical experiment*. As a simple illustration, consider the problem of multiplying two numbers x and y. A 'digital' method would be to look up an explicit multiplication table. An alternative, however, would be to take two physical quantities (such as the distances along the graduated scales of the once-familiar slide rule) proportional to the logarithms of x and y, and perform the physical experiment of combining them so that their physical sum determines the number read off the 'product' scale. In this so-called 'analogue' method, the actual outcome only approximates to the ideal, so it is unsuitable for precise symbolic calculations; but there are many situations, both in automata and in the human brain, where internal physical experimentation of this kind plays a crucial and effective part in determinig what happens next. The point is that, as with all appeals to the real world, the outcome of such processes have a degree of *contingency*, so that the functioning of the corresponding mechanism is irreducible to the making of selections uniquely determined by rule. (The digital computer can be regarded as a special case in which this contingency is made negligible by arranging that all physically-likely outcomes of a given symbolic situation are symbolically equivalent.)

Now the brain, considered as an information-processing mechanism, combines processes of both types. Neurones of the CNS are metastable elements which interact in a cooperative fashion, such that at a given time there is for each network a vast array of 'conditional probabilities' that *if* event A occurs *and if* B occurs within so many milliseconds, *then* the network will go over from state *i* to state *j* with transition-probability *Pij*, and so on. These states may form a discrete set; but the physical quantities that govern the *Pij* (ionic concentrations, proximity of synaptic terminals and the like) are generally continuously variable, and interact so delicately that no predictions based on the physical measurement of parameters could keep track of the outcome, even in principle, for a significant time. The best that physics could do, even with full information, would be to specify at any time a vast 'transition-probability-matrix' (TPM) indicating the conditional probabilities of various outcomes in various circumstances. Brain activity has to be thought of as a stochastic process, to whose actual time-course no explicit program of determinate rules could be equivalent.[13]

The consequences for the hypothesis of our embodiment as conscious agents are far-reaching. In the first place, since all that characterizes us as individuals (our memories, dispositions, skills etc.) depends *ex hypothesi* on the detailed physical structure upon and in which the aforementioned 'internal experimentation' is carried out, and on which each twist and turn of our experience leaves its mark in a widely distributed and largely implicit physical form, it is impossible even in principle to divorce our personal identity from the specific brain-matter in which we are embodied, in the easy way in which the identity of an equation can be divorced from its specific embodiment in a digital computer. I am the individual whose brain-matter went through this and that specific course as a result of these and those specific internal experiments in and upon it. You are another individual whose brain-matter has traced a differ-

ent world-line. Nobody with a different world-line can be identical with you or me, for his physical embodiment is unique to him, a unique theatre of physical experimentation, such that even if he had been equipped instantaneously (in the manner of some science-fiction philosophers) with the best possible physical replica of your brain, his would manifest immediate and increasing divergences from the path taken by yours. I do not deny outright the notion that in the science-fiction world the 'replica' might share many of your memories or mine; and it is perhaps arguable that he might be allowed to share some of your retrospective obligations in these bizarre circumstances. What seems quite clear is that (if the replica is a conscious being at all) he has no rational claim to *be* you or me, whether or not you or I are killed (for his convenience) as soon as he is brought into being. Much of the extensive current argumentation about such possibilities[14] is vitiated by failure to recognize the differences between an imaginary 'brain' that is functionally equivalent to a program of easily replicable rules, and the real brain in its physical particularity as a theatre of continual appeals to local physical reality, the outcomes of which play an essential part in shaping both the course and the character of our embodied agency. In short, for us as embodied persons, "matter matters".

What I am arguing for is not materialism, but what I would call "Duality without Dualism".[15] The data of our conscious experience demand to be reckoned with as categorically distinct from, and ontologically prior to, all data about our brain activity. However tight the correlations we may discern between them, their duality seems at least as undeniable as the duality between the mathematical and the physical aspects of what goes on in a computer solving an equation. By the same token, there seems to be as little need for the thought-model of interactionist dualism, with its talk of causal traffic between two 'worlds', in the one case as in the other. The relationship between embodied and embodiment is *more intimate* than that between cause and effect.[2] Unlike Searle,[16] I would argue[3, 5] that to attribute 'thinking' or other intentional activities to either brains or machines is to confuse categorical levels. It is not brains, but persons, that think, hope, fear and the like. To insist (as I do) that the physical constitution of the brain is important for personal identity does not at all imply that it makes sense to attribute mental activity to the physical brain.

Secondly, if the structure of the brain is thought of as constraining and enabling our behaviour in anything like the sense in which the program shapes the behaviour of a computer, we must recognize that (in contrast to a digital computer program) much of its contribution takes the form of "inclination without necessitation" (another expression used by Chisholm (1982, pp. 33–35), from a different standpoint). For reasons given below I think we should resist the temptation to take this stochastic aspect of brain processes as the physical basis of freewill;[2] but it does at least dispose of the idea that mechanistic theories of brain function can find no scope for genuine creativity.

Thirdly, to conceive of our embodiment in these terms disposes of any idea that mechanistic explanation of brain function would leave no room for what Wittgenstein [17] termed the 'ineffable' (that which cannot be put into words but can only be 'manifested'), or for what Polanyi[18] called 'tacit knowledge'. That "we know more than we can tell" (Polanyi) would follow directly from the hypothesis that we are embodied in a mechanism in which information is represented partly in the implicit form of physical constraints on the relative probabilities of transitions in the course of internal experimentation. The knowledge so embodied would be tacit, in the sense that its *possessor* would have no words for it, though it would be immediately accessible to him in the course of conscious agency, in the way that knowledge how to swallow, or how to ride a bicycle, can be.[19] By the same token, mechanistic explanation of brain function on these lines would positively encourage expectations that the possessor could sometimes be aware of 'unutterables' of his experience which might nevertheless (in Wittgenstein's terms) be 'shown', through the medium of poetic or artistic expression.[20]

The stochastic nature of brain activity has a further consequence worth noting. Our normal thinking is frequently enlivened by what we call 'unbidden' events. 'It occurred to me that X', we say, or 'I was struck by Y'. Thinking, in other words, is a compound of 'doing' and 'suffering'. Now on our hypothesis of embodiment, the conditional probabilities of such unbidden

transitions presumably depend on the microphysical details of specific structures (synapses or the like) in the thinker's brain. There is abundant and tragic evidence from the effects of drug abuse that gross chemical factors can modulate these conditional probabilities. What is less often recognized is that in an associative mechanism on these stochastic lines, even the speculative 'trial running' of ideas can be expected in general to change significantly the relative probabilities of future spontaneous transitions under *given* conditions. In other words, to be embodied in such a mechanism means that even speculative questioning is not a neutral activity, but can play a significant part, willy nilly, in determining what is likely to 'occur to' the questioner for the whole of his future life. The moral for the philosopher may be worth pondering!

A prime cost of mechanistic embodiment is that there can be no gain (or loss) of knowledge without a change in the physical world—specifically, a change in whatever cerebral structure has to embody the information in question. Popper,[21] in a classic paper of 1951, argued that even in a physically determinate universe this must set restrictions on the predictability of the universe for its inhabitants.

It has however a further consequence, which I have followed up elsewhere.[22–24] Mechanistic embodiment entails certain *relativistic restrictions* on the validity of state-descriptions of embodied persons. A complete and uptodate state-description S of the immediate future of the embodiment of an agent A, assuming it could be produced by a non-interfering observer O, would be valid-for-O but systematically invalid-for-A. This says more than that A could not know or believe S: it means that *no uniquely determinative specification exists*, even unknown to A, which has an equally valid claim to the assent of A and of O. Instead, we must recognize that what A would be correct to believe, and in error to disbelieve, about his immediate future is systematically different from what O would be correct to believe about it. Even if, for O, every detail of A's future had a fully determinate specification, it would and could have no such complete specification with an unconditional claim to A's assent, until after A's cognitive agency had determined relevant details of that future. I will not here retrace the consequences of this 'logical indeterminacy' for the freewill debate.[23, 26] I want only to emphasize the extent to which some standard non-relativistic forms of argument can be undermined by it. In particular, the question whether a specific future event is 'inevitable' is *not properly formulated* until we say *for whom* the inevitability is to be assessed. Even in a physically-determinate universe an event inevitable-for-O may not be at all inevitable-for-A, if A's future cognitive activity is one of the factors that will determine its form.[24]

Recognition of this relativistic principle does nothing to question the validity of standard logic, since nothing but standard logic is necessary to demonstrate it. All it means is that if we try to use standard logic to determine what cognitive agents would be correct to believe about themselves, we are not entitled to presuppose (as conventional logical arguments tend to do) that if anyone would be correct to believe proposition p, then everyone would be. It follows that the luxury of being able to indicate the truth-value of p universally by a single-valued scalar (true or false, 1 or 0) may in some such cases be denied us. In this strict sense the case of cognitive agency has about it an irreducible element of *mystery*, whatever the degree of physical explicability we attach to the workings of its embodiment.[25, 26]

What then of the common notion that the hypothesis of mechanistic embodiment entails rejection of all religious assessments of our spiritual significance and destiny? Much argumentation on these lines is based on the mistaken inferences we have already considered; but it could be thought that by emphasizing the link between our personal identity and our physical embodiment we have made it more difficult to take seriously the concept of 'eternal life' as found for example in biblical Christianity. This is not the place for a full discussion;[2, 26] but the apparent difficulty would arise only if the 'eternal life' in question were a mere disembodied persistence of the individual within our present space-time. For the Christian religion, which maintains that a personal Creator is the 'ground of being' of the physical world as well as of the persons embodied in it, the emphasis is not on mere persistence but on something even more mysterious—'resurrection' in a new embodiment, and moreover in a 'new creation'—a

different space-time. On the hypothesis we have been examining, the continuity required in the embodiment would not be that of the material elements of the brain structure, but of its (to us unspecifiable) transition-probability matrix (TPM). Irrespective of one's individual convictions on these matters, there are clearly no rational scientific objections to the doctrine that the Creator who sustains our present TPM from one day to the next can ensure all the continuity required in the resurrection, if He will.

In conclusion it may be useful to take stock. I have argued that we must abandon the Procrustean and physiologically unrealistic image of our cerebral embodiment as a species of rule-determined symbol-manipulator, and recognize the functional significance of the brain as a theatre of internal physical experimentation, bearing the traces of its whole history in the transition-probability matrix implicit in its physical microstructure. The neuroscientific working hypothesis that we are each embodied in such a mechanism, I have argued, would *not* entail the reducibility of all our behaviour to rule-following. It would *not* entail any denial of our ability, by our thinking, valuing and deciding, to determine our own future. It would *not* imply that I could be 'teletransported' from A to B merely by having someone programme a mechanism at B to function in an identical manner to my brain. It would not even imply that a mechanism so programmed could be guaranteed to embody any conscious experience whatever. It would raise no rational obstacle, however, to the possibility of 'eternal life' as adumbrated in the Christian religion.

I have further argued that mechanistic embodiment on these lines entails no denial of genuine creativity, nor of the reality and importance of tacit human knowledge inexpressible in words. Its main 'cost' is that, if for us there can be no coming-to-know without a change in the physical world, then no complete (present or future-tense) specification of the physical world can exist with an unconditional claim to the assent of all cognitive agents. Instead, in the case of cognitive agency we must be content to recognize a relativistic situation, in which it is not necessarily the case that if O would be correct to believe p and in error to disbelieve p, then everyone else would be correct to believe p and in error to disbelieve p. In particular, for an agent A contemplating a choice, the immediate future of his cognitive system is not just unknown but logically *indeterminate* until he makes up his mind, even if his brain were as physically determinate as clockwork. Thus the costs of mechanistic embodiment do *not* include any implication that the future outcome of a choice is already inevitable for the mechanistically embodied chooser.

It will, I hope, be clear that in this whole discussion I have had no intention of urging a belief in mechanistic determinism, or even in computational models of the brain. My aim has rather been to loosen up a knotted skein of improperly linked inferences in order that the hypothesis of our mechanistic embodiment may be disentangled and considered on its merits. The question whether all human brain activity has a mechanistic explanation, I suggest, is one we can peacefully leave open for future investigation, no matter how high a view we take of man's power of decision and its moral and religious significance.

ENDNOTES

* The numbers in round brackets refer to the corresponding endnotes.

[1] MacKay, D. M., "Mindlike Behaviour in Artefacts", *Brit. J. Phil. of Sci.*, Vol. 2 (1951), pp. 105–121. Also Vol. 3(1953), pp. 352–353. "Mentality in Machines" *Proc. Aristot. Soc. Suppt.* Vol. 26 (1952), pp. 61–86.

[2] MacKay, D. M., "Man as a Mechanism", *Faith and Thought* Vol. 91 (1960), pp. 145–157. Also revised in D. M. MacKay (ed.), *Christianity in a Mechanistic Universe* (London 1965).

[3] MacKay, D. M., "The Use of Behavioural Language to Refer to Mechanical Processes", *Brit. J. Phil. of Sci.* Vol. 13 (1962), pp. 89–103. Also in F. J. Crosson (ed.), *Human and Artificial Intelligence* (New York 1971).

[4] MacKay, D. M., "The Interdependence of Mind and Brain", *Neuroscience* Vol. 5 (1980), pp. 1389–1391.

[5] MacKay, D. M., "Mind Talk and Brain Talk", in: M. S. Gazzaniga (ed.), *Handbook of Cognitive Neuroscience* (New York 1983), pp. 293–317.

[6] MacKay, D. M., "Complementarity II", *Aristotelian Soc. Suppt.* Vol. 32 (1958), pp. 105–122. "Complementarity in Scientific and Theological Thinking", *Zygon* Vol. 9 (1974), pp. 225-244.

[7] Sperry, R. W., *Science and Moral Priority*, (Oxford 1983).

[8] While accepting that at the physical level "the laws of causation are nowhere broken or open (excepting perhaps in quantum-level indeterminacy which is here irrelevant)" Sperry still claims to "counter and refute . . . the classic physicalist assumption of a purely physical determinacy of the central nervous system" (Sperry 1983) pp. 81, 89). If by "purely" here he only means "exclusively" I would agree (see notes 1–5); but the confusion that can be caused by Sperry's choice of terms is brought out if we imagine a parallel argument in the case of our digital computer.

[9] Chisholm, Roderick M., "Human Freedom and the Self", in Gary Watson (ed.), *Free Will* (Oxford 1982), p. 28.

[10] MacKay, D. M., "Machines, Brains and Persons" *Zygon*, (1985, in press).

[11] Turing, A. M., "Computing Machinery and Intelligence", *Mind* Vol. 59 (1960), pp. 433–60.

[12] The term 'personality' has an unfortunate ambiguity. (a) We speak of the 'personality' of a character in a play or a novel, meaning (roughly) the dispositional features that distinguish that character psychologically, without any implication that he exists as a person. But when we ask whether we can rightly attribute 'personality' to a robot, we may be raising (b) the ontological question whether the robot *is* a person (a centre of conscious awareness), as distinct from merely behaving *like* a person.

I must plead guilty to some inconsistency in past uses of the term. In "The use of behavioural language to refer to mechanical processes" (note 3) I argued that a suitable artefact might have a personality in sense (a). In "A mind's eye view of the brain" (note 25) I used the term in sense (b) and denied that behavioural criteria were sufficient to settle the *ontological* question of personality. In each case I would still affirm what I meant, which I hope was clear from the context.

[13] MacKay, D. M., "In Search of Basic Symbols", in: H. von Foerster (ed.), *Proc. 8th Conf. on Cybernetics* (New York 1951), pp. 181–221; "From Mechanism to Mind", *Trans. Vict. Inst.* Vol. 85 (1953), pp. 17–32; "Operational Aspects of Some Fundamental Concepts of Human Communication" *synthese* Vol. 9 (1954), pp. 182–198; "Operational Aspects of Intellect" *Mechanization of Thought Processes* (Her Majesty's Stationery Office 1959), pp. 37–52; "Digits and Analogues", in: H. E. von Gierke, W. D. Keidel and H. L. Ostreicher (eds.) *Principles and Practice of Bionics* (London 1970), pp. 457–466; see also note 5.

[14] See for example Derek Parfit, *Reasons and Persons* (Oxford 1984).

[15] MacKay, D. M., "Ourselves and our Brains: Duality without Dualism", *Psychoneuroendocrinology* Vol. 7 (1982), pp. 285–294. See also MacKay (1980), note 26.

[16] Searle, John R., "Minds, Brains and Programs", *The Behavioral and Brain Sciences* Vol. 3 (1980), pp. 417–458.

[17] TLP 6.432: "There are indeed things that cannot be put into words. They *make themselves manifest*".

[18] Polanyi, M., *The Tacit Dimension* (London 1967).

[19] MacKay, D. M., "The Mechanics of Tacit Knowing", *I. E. E. E. Trans. on Systems, Man and Cybernetics* SMC-4 No. 1 (1974), pp. 94–95.

[20] Witness Wittgenstein's comment in a letter of 9th April 1917 to Engelmann: "The poem by Uhland is really magnificent. And this is how it is: if only you do not try to utter what is unutterable then *nothing* gets lost. But the unutterable will be–unutterably–*contained* in what has been uttered!" See P. Engelmann: *Letters from Ludwig Wittgenstein* (Oxford 1967), p. 7.

[21] Popper, K. R., "Indeterminism in Quantum Physics and in Classical Physics", *Brit. J. for Phil. of Sci* Vol. 1 (1950), pp. 117–133 and 173–195.

[22] MacKay, D. M., "Scientific Beliefs about Oneself", in: G. N. A. Vesey (ed.), *The Proper Study* (1971), pp. 48–63.

[23] MacKay, D. M., "On the Logical Indeterminacy of a Free Choice", *Proc. XIIth Int. Congress of Philosophy* Vol. 3 (1958), 249–256; and expanded version in *Mind* Vol. 69 (1960), pp. 31–40. *Freedom of Action in a Mechanistic Universe* (Eddington Lecture), (London, New York 1967), reprinted in M. S. Gazzaniga and E. P. Lovejoy (eds.), *Good Readings in Psychology* (New York 1971), pp. 121–138.

[24] MacKay, D. M., "What Determines My Choice?", in: P.A. Buser and A. Rougeul-Buser (eds.), *Cerebral Correlates of Conscious Experience* (Amsterdam, New York, Oxford 1978), pp. 335–346.

[25] MacKay, D. M., "On Comparing the Brain with Machines" *The Advancement of Science* Vol. 40 (1954), pp. 402–406. Also *American Scientist* Vol. 42 (1954), pp. 261–268; and "A Mind's Eye View of the Brain", in: Norbert Weiner and J. P. Schade (eds.), *Cybernetics of the Nervous System, Progress in Brain Research* Vol. 17 (1965), pp. 321–332; "Selves and Brains", *Neuroscience* Vol. 3 (1978), pp. 599–606.

[26] MacKay, D. M., *Human Science and Human Dignity* (London, Downers Grove, Ill., 1979). *Brains, Machines and Persons* (London, Grand Rapids, 1980).

* * *

GEHIRNTÄTIGKEIT UND BEWUSSTSEIN

Franz Seitelberger
Universität Wien

Das Wort Bewußtsein (B) wird in diesen Ausführungen als Bezeichnung einer Tätigkeitsweise des Gehirns, nämlich des bewußten Verhaltens, verwendet. Das zentrale Anliegen der Hirnforschung besteht in der Demonstrierbarkeit der psychischen Phänomene als Hirnleistungen. Die bisherigen Ergebnisse dieses Unterfangens legen nahe, daß es keine definierbaren morphologischen noch physiologischen Substrate *des* Bewußtseins gibt, sondern daß das B eine in der strukturellen Organisation der neurobiologischen Informationsverarbeitung hervortretende Disposition, und zwar jene der subjektiven Erscheinungsweise der individuellen Verhaltenssteuerung darstellt. Das B gibt es somit nur im Sinn der bewußten Beziehung zwischen Individuum und Umwelt in Form der bewußten höheren Hirnleistungen. Zu diesen zählen unter anderem Gedächtnis, Gegenstandswahrnehmung, Sprache und Denken. Diese Worte stehen aber nicht für einfache Leistungen, sondern für noch unbekannte Leistungskomplexe und hierarchisch organisierte Programme, für deren Durchführung die Großrinde unentbehrlich ist.

Man kann in der *Großrinde* 3 Gruppen von Rindengebieten hinsichtlich ihrer funktionalen Eigenart unterscheiden: Rindenregionen, in denen bestimmte Teile des Informationseinflusses ein- bzw. austreten, z.B. die primären Sinnesfelder (Sehregion etc.) oder die motorische Region: *Modale Rindenfelder*.

Gemäß dem, was über die sogenannte modulare Organisation der GHR bekannt ist, bedeutet die Aussage, daß in solchen Regionen bestimmte Funktionen *lokalisiert* seien, nicht eine funktionelle Spezität des betreffenden Hirnteiles, sondern die lokale Organisation des betreffenden Informationsinputs oder -outputs. Die Arbeitsweise der Hirnrinde ist prinzipiell in jeder Region die gleiche; die Arbeitsergebnisse aber sind vom bearbeiteten Material abhängig: Sinneserfahrungen, Bewegungen, etc. Die Konzentration der äußeren Informationskanäle (nervöse Leistungsbahnen) in solchen Regionen bedingt es auch, daß ihre Läsionen ganz bestimmte und konstante Funktionsfälle zur Folge haben, z.B. Blindheit oder Lähmung.

Anders verhält es sich mit den *intermodalen Rindenfeldern*, die sich zwischen den modalen ausspannen und diese in ihrer Ausdehnung beim Menschen weit übertreffen. Sie erhalten als Input die bereits von den modalen Feldern bearbeitete Information, also sekundäre, tertiäre etc. Verarbeitungsprodukte und zwar konvergierend von mehreren Modalfeldern sowie anderen Rinden- und Hirngebieten. Ihre Arbeit besteht in der integrativen Aufarbeitung großer Informationsmengen verschiedener Herkunft, sei es zum Aufbau der einheitlichen Wahrnehmungswelt oder zur optimalen Bewegungssteuerung. Die Läsion solcher Rindengebiete verursacht keine groben Defekte, man bezeichnet sie daher als „stumme“ Regionen, sondern Beeinträchtigungen der Funktionshöhe. Die repetitive und distributive Bearbeitung jeden Inputs durch den Modulapparat der GHR erreicht im Menschengehirn dadurch eine maximale Stufe, daß große Gebiete vor allem des Stirn- und Schläfenlappens die Arbeitsergebnisse aller anderen Hirnteile zusammenzuführen und nochmals analysieren. Man kann umschreibend sagen, daß in diesen Regionen die Integration aller dem Gehirn zugehender Information in Beziehung auf die Gesamtsituation des Individuums mit Extrapolation in die Zukunft vorgenommen wird. Das betrifft nicht nur die Information aus der Außenwelt, sondern auch die aus dem eigenen Körper einschließlich der Befindens- und Triebsphäre.

Aus dieser „Information über Information“ erwachsen offenbar die den eigentlich menschlichen Hirnleistungen korrelierten Fähigkeiten des Lernens, Denkens, Planens und Handelns,

die man auch als *supramodale Funktionen* bezeichnet. Das entscheidende Moment dieser Leistungsstufe liegt in der Erschließung einer neuen Möglichkeit des Umfangs mit symbolischen Produkten der Hirntätigkeit, nämlich in der Behandlung der Abstraktionen, in denen für die Verhaltenssteuerung durch das Gehirn Wirklichkeitsäquivalente repräsentiert sind. Abstraktion ist zwar ein Grundzug der kodierten Datenverarbeitung überhaupt und gehört z.B. in Form der „feature extraction" zur Voraussetzung der einfachsten Wahrnehmung; in der supramodalen Stufe erreicht die Abstraktionsfähigkeit aber erst die höchste Form begrifflicher Kristallisation von innerer und äußerer Wirklichkeit. Sie manifestiert sich in der Sprache und ist allen höheren Hirnleistungen inhärent. Sprachfähigkeit bedeutet, daß die Integrationsprodukte höherer Ordnung, wie sie z.B. dem einheitlichen Wahrnehmungsgegenstand oder einem Bewegungsentwurf zugrunde liegen, Produkte, die auf dem Niveau der Gehirntätigkeit als komplexe raumzeitliche Erregungsmuster vorzustellen sind, als selbständige Einheiten identifiziert und klassifiziert werden, z.B. als ein „Baum", und in solcher Gestalt einer einfachen nicht-neurophysiologischen Neukodierung unterworfen werden können, nämlich der im geschichtlichen kulturellen Prozeß geschaffenen menschlichen *Sprache*. Der „Baum" ist also ein Konstrukt, der aus vielen Informationsquellen gespeisten integrierenden Hirntätigkeiten, ein Konstrukt, das in der lautlichen Kodierung als „Baum" wieder einen Input darstellt, der in den Modulapparat der GHR leicht eingespeist und bearbeitet werden kann, wobei die enorme Beziehungsfülle des Konstruktes in der knappen Form „Baum" erhalten bleibt und die Neubearbeitung einer Potenzierung der Resultate erbringen kann. Diese von der Sprachfähigkeit getragene Symbolbildung geht aber über die Konstruktion abstrakter Modelle von realer Wirklichkeit noch hinaus: In Verbindung mit der Lernfähigkeit und dem Gedächtnis gewinnt die integrative Hirntätigkeit Unabhängigkeit von konkreter Erfahrung, sie erweist sich quasi als ein unspezifisches Sinnesorgan, das seine Objekte selbst hervorbringt: Das ist die *Vorstellungskraft*. Mit ihr ist es nun möglich, nicht nur die ankommenden Informationen über aktuelle Lebenssituationen optimal auszuschöpfen, sondern auch die mittels der Vorstellungskraft in Eigentätigkeit erstellten Modelle von möglichen Situationen und zukünftigen Ereignisfolgen auszuwerten. Man bezeichnet den durch die supramodale Rindenfunktion gewonnenen Erfahrungshorizont, der den durch die Sinne vermittelten Zustand des aktuellen Lebensraumes, zugleich aber auch die Situations-relevanten, mittels des Gedächtnisses und der Vorstellungskraft gebildeten Wirklichkeitsmodelle mit Prospekten in die Zukunft umfaßt, also ein für die individuelle Existenz relevantes Weltbild vermittelt, auch als *Intelligenz* und ihre Betätigung als *Denken*.

Bei dieser Leistungsschilderung der Großhirnregionen muß die Leistungsdifferenzierung zwischen den beiden Hemisphären, also die sog. *Hemisphärenspezialisierung* berührt werden. Sie ist seit langem in Form von *Dominanz* der linken Hemisphäre des Rechtshänders für den Gebrauch der Hand und für die Sprachfähigkeit bekannt.

Sehr vereinfacht und verallgemeinert ausgedrückt, kann man der linken Hemisphäre des Rechtshänders eine analytisch-logische Funktion über der Zeitachse und der rechten Hemisphäre eine synthetisch-perzeptive Funktion über den Raumkoordinaten zusprechen. Wichtige Aufschlüsse über die Beziehung zwischen Gehirn und bewußten Leistungen erbrachten die sogenannten split brain-Experimente. Patienten mit durchtrenntem oder fehlendem Hirnbalken können auch mit der isoliert arbeitenden rechten Hemisphäre korrekt erkennen und mittels der von ihr gesteuerten linken Hand entsprechend agieren. Diese Leistungen haben aber für sich allein nicht den Charakter bewußter Wahrnehmungen bzw. Handlungen und sind daher nicht sprachfähig. Nichtsdestoweniger ist der Beitrag der rechten Hemisphäre zum bewußten Verhalten unentbehrlich und konstitutiv. Alle höheren Hirnleistungen kommen durch Zusammenwirken beider Hemisphären zustande.

Man erkennt, daß es sich bei der verhaltenssteuernden Gesamtfunktion des Gehirns nicht darum handelt, ein wissenschaftlich exaktes Abbild der Umwelt als Aktionsbasis zu liefern, sondern eine umfassend informierte und abgestimmte *Verhaltensbereitschaft*, von der aus die Handlungsplanung in allen relevanten Umständen bestimmt, bzw. begrenzt werden kann.

Die Einheit des organisierenden Hirnsystems entspricht der Einheit der bewußt erlebten

Welt und stellt den durch die Sinnesmannigfaltigkeit gestalteten Erfolgsraum der individuellen Aktionen dar.

Von hier aus können wir noch einen kurzen Blick auf die *metamodale Ebene der Hirntätigkeit* richten, auf der ihre Leistungsprodukte sich selbst zum Objekt werden, also eine Reflexionsebene, um uns einen Begriff davon zu machen, was *Bewußtsein* ist oder sein kann: jedenfalls keine eigene Funktion bestimmter Hirnapparate.

In einer die bisherigen Feststellungen erweiternden Hypothese könnte man sagen: Das B ist die durch die strukturelle Organisation der Informationsverarbeitung bedingte, subjektive Erscheinungsweise unserer Verhaltenssteuerung. Die Einheit unserer Existenz ist uns paradoxerweise nur in dualer Weise erfahrbar: in der Wahrnehmung der Welt und im Erlebnis des Ichs. Sobald durch die vervollkommnete Ansicht der Gesamtsituation und durch die autonome Produktion von beliebigen Verhaltensmöglichkeiten durch die Vorstellungskraft dem Individuum im raumzeitlichen Weltgeschehen eine Mehrzahl von Alternativen des Verhaltens mit Zukunftsprojektion zur Verfügung stehen, erfolgen bestimmte optimierende Entscheidungen im Medium des B. Das B eröffnet die für diese Verhaltensebene erforderliche zusätzliche neue logische Dimension (MacKay) bzw. liefert dazu einen eigenen anderen Code (Pribram). B ist also eine mit der Höhe und Komplexität der Informationsverarbeitung von Lebewesen immanent verbundene *Qualität*. Im B erlebt das Individuum unmittelbar seine relevante Weltbeziehung in Hinblick auf die ihm möglichen und daher anheimgestellten Entscheidungen. Das elementare Verhalten der Lebewesen zur Umwelt in Form von Reiz und Reaktion erhebt sich damit beim Gehirnträger auf die Ebene von B und Wahl, beim Menschen schließlich zur Ebene von Wissen und Wollen. Die Tatsachen des B sind somit eigengesetzliche Äußerungen des einen ganzen Individuums, in denen die Totalität seiner Organismus-Umwelt-Beziehung subjektiv repräsentiert wird. Beim Menschen erreicht dieses bewußte Inbild der Wirklichkeit die äußeren Dimensionen der Welt und der Zukunft, sowie die innere Dimension des Selbst, der autonomen menschlichen Person.

LITERATUR

Creutzfeldt, O. D., „Bewußtsein und Selbstbewußtsein als Problem der Neurophysiologie", *Universitas* 36, S. 467–475 (1981).

MacKay, D. M., *Brains, Machines and Persons*. (London 1980).

Pribram, K., „Holographic Brain", in: P. Weintraub (Ed.), *The Omni Interviews*. (New York 1983) S. 133–151.

Seitelberger, F., „Die Raum-Zeit-Struktur der menschlichen Erlebniswelt als Problem der Hirnforschung", in: *Studien zur Klassifikation*, Bd. 9 (Frankfurt 1980), S. 178–196.

Seitelberger, F., „Die Evolution der Erkenntnis. Leistungspotenzen und Leistungsprodukte des menschlichen Gehirns", in: *Erkenntnis- und Wissenschaftstheorie. Akten des 7. Internationalen Wittgenstein-Symposiums 22.–29.8.1982* (Kirchberg/Wechsel), S. 174–184.

Seitelberger, F., „Neurobiological Aspects of Intelligence," in: Franz Wuketits (Ed.), *Concepts and Approaches in Evolutionary Epistemology* (Dordrecht-Boston-Lancaster 1984), S. 123–148.

Sperry, R. W., Gazzaniga, M. S. and Bogen, J. E., „Interhemisphere Deconnection," in: P. J. Vinken & A. W. Bruyn (Eds.), *Handbook of Clinical Neurology*, Vol. 4 (Amsterdam-New York-Oxford 1978), S. 273–290.

Szentagothai, J., The local neuronal apparatus of the cerebral cortex. In: P. A. Buser and A. Rougeul-Buser (Eds.), *Cerebral correlates of conscious experience* (North-Holland, Amsterdam-New York-Oxford 1978), S. 131–138.

Wuketits, F. M., „Das Problem des Bewußtseins aus biologischer Sicht," *Österreichische Ärztezeitung* 35/1 (1980), S. 24–26.

* * *

EMOTION UND SPRACHE – EIN PSYCHOLOGISCHER ANSATZ

Peter Vitouch
Universität Wien

Der vorliegende Beitrag möchte als ein Denkanstoß gewertet werden, experimentalpsychologisch gewonnene Erkenntnisse mit philosophischem Instrumentarium auf ihre Haltbarkeit zu überprüfen. Es geht dabei nicht um eine wechselseitige Methodenkritik, sondern um die Darstellung der Möglichkeit, psychologische Theorien, die auf empirischen Daten fußen, in einen philosophischen Zusammenhang zu stellen.

Es geht um die Entwicklung der aktuellen psychologischen Emotionstheorien; also der Theorien, welche die Entstehungsweise eines Gefühls erklären wollen. Paradoxerweise ist der Terminus Emotion sehr aussagekräftig für die Allgemeinheit, für den experimentell orientierten Forscher ist Gefühl dagegen eine schwer definierbare und objektivierbare Größe.

Aristoteles unterschied als erster zwischen einer physiologischen und psychologischen Komponente der Emotion, die er einerseits als „Materie", andererseits als „Form und Idee" bezeichnete. Für die moderne Psychologie erwiesen sich die Erkenntnisse aus der Gehirnphysiologie und Neuropsychologie als bedeutsam, weil damit Phänomene des Erlebens, Denkens und Fühlens näher beschrieben werden konnten. Es wurde immer klarer, daß die Funktionen peripherer Systeme und ihre Interaktionen in der Hauptsache von Steuer- und Regelzentren des Gehirns kontrolliert werden (vegetatives Nervensystem). Ebenso besteht heute kein Zweifel darüber, daß psychologische Phänomene zentralnervöse Grundlagen besitzen, daß also die Funktion unseres Gehirns verantwortlich dafür ist.

Die Gehirnphysiologie machte sich an eine „Vermessung" unseres Gehirns, entdeckte motorische und sensorische Zentren, entwickelte eine „Kartographie" der Gehirnfunktionen. Nahezu alle Fähigkeiten des Menschen wurden bestimmten „Zentren" zugeordnet und demnach wurden auch den Gefühlen sogenannte „Emotionszentren" zugeordnet. Es zeigte sich jedoch, daß diese Auffassung eine unzulässige Trivialisierung der Funktionsweise unseres Gehirns darstellte. Sicherlich gibt es Zellformationen in unserem Gehirn, in denen zum Beispiel Gefühlsreaktionen gespeichert sind (wir können uns an einmal erlebte Gefühle erinnern). Die Existenz derartiger Zentren reicht aber sicher nicht aus, den Vorgang der Entstehung eines Gefühls zu repräsentieren. Ein Vorgang, an dem – wie man heute weiß – nahezu das ganze Gehirn in differenzierter Interaktion beteiligt ist.

Die nähere Untersuchung der als „Sitz des Gefühls" postulierten Gehirnareale führte jedoch zu interessanten Erkenntnissen der Neuropsychologie auf dem Gebiet der sogenannten „Aktivierungsforschung".

Was ist Aktivierung? Dem Zustand jedes Menschen kann in jedem Moment seines Lebens ein Grad auf einer gedachten Skala zugeordnet werden, die von tiefem Schlaf bis zu höchster Angespanntheit (z.B. Panik) reicht. Diese Abstufungen werden als Grade „innerer psychischer Entspanntheit bzw. Erregung" erlebt. Wir befinden uns also zu jeder Zeit auf einem bestimmten „psychischen Aktivierungsniveau", das sich aufgrund unterschiedlicher Umweltbedingungen mehr oder weniger stark ändern kann. Berlyne (1960) spricht vom sogenannten „Aktivierungstonus", das ist der Grad einer mittleren Aktiviertheit, den jeder Mensch zu erhalten versucht, weil er als angenehm empfunden wird.

Lange Zeit wurde die Formatio reticularis (ein Nervennetzwerk im Hirnstamm) als alleinverantwortlich für die Aufrechterhaltung des Aktivierungsniveaus angesehen. Magoun (1948) entdeckte, daß die Formatio reticularis eine „Weckwirkung" auf die Großhirnrinde ausübt und dadurch von wesentlicher Bedeutung für die jeweilige Bewußtseinshelligkeit ist.

Beim Studium der Formatio reticularis durch Olds und Milner (1958) verfehlte eine

Tiefenelektrode das Ziel und wurde in das Septum eines Versuchtstieres implantiert. Es zeigte sich, daß die Tiere die elektrische Reizung dieser Region jedem anderen Reiz vorziehen. Schließlich ermöglichten die Versuchsleiter dem Tier, sich selbst zu reizen, indem das Tier den Stromkreis durch Hebeldruck selbst schließen konnte. Das Ergebnis war, daß die Tiere nun häufig bis zu 5000mal und mehr pro Stunde drückten, und dies bis zur völligen Erschöpfung. Selbst wenn dem Tier nach langer Hungerperiode Nahrung angeboten wird, wenn es Durst hat oder sexuelle Reize es ablenken sollen, wird die Selbstreizung jeder anderen Reaktion vorgezogen. Die Reizung tiefer gelegener Strukturen des Mittelhirns, im sogenannten periventrikulären System, hatte gegenteiligen Effekt. Die Tiere versuchten mit allen Mitteln, jede Art von elektrischer Reizung dieser Hirnteile zu vermeiden. Olds definierte diese Strukturen als „pleasure centers“ (Belohnungszentren) und Bestrafungs- oder Aversionszentren.

Damit sind wir schon auf der Spur der Gefühle. Rohracher (1965) sagt: „Die Dimension Lust-Unlust, oder in neuerer Zeit die Polarität Verstärkung versus Aversion, werden als die bestimmenden Merkmale von Gefühlen angesehen.“ Jedenfalls gibt es keine Gefühle, die wir nicht zumindest in geringem Maße angenehm oder unangenehm empfinden.

Berlyne (1969) konnte in Experimenten nachweisen, daß offenbar die Intensitätskomponente (in Bezug auf Aktivierung) in Zusammenhang mit Lust-Unlust-Empfindungen steht.

Diese Ergebnisse stimmen mit der Alltagserfahrung überein. Jeder angenehme Reiz ist nur solange angenehm, solange er nicht zu intensiv wird. Der von Berlyne definierte Aktivierungstonus wird sich also auf einem mittleren Niveau einpendeln, um als angenehm erlebt zu werden.

Der nächste wesentliche Punkt, der menschliches Verhalten erklären hilft, ist die Einführung des Terminus ‚Aktivierungsspitzen‘ Diese sprunghafte Veränderung des Aktivierungstonus ist das Resultat von Reizen aus der Umwelt, auf die ein Individuum reagiert. Als angenehm oder belohnend wird der – in mehr oder weniger kurzer Zeit – darauf folgende Rückgang des Aktivierungsniveaus erlebt.

In zahlreichen Experimenten der Lernpsychologie wurde bewiesen, daß ein Organismus auch Unangenehmes anstrebt (in diesem Fall das Abweichen von Aktivierungstonus), wenn dieser unangenehme Reiz ein Signal für darauffolgende Belohnung darstellt. Man strebt also Situationen an, welche die Aktivierung kurzzeitig anheben, unter der Voraussetzung, daß nach absehbarer Zeit eine Rückkehr zum Aktivierungstonus stattfindet.

Trotz der Plausibilität, die diese Theorien entwickeln, haben sie den gewichtigen Nachteil, die Funktion des Cortex (der Gehirnrinde) nicht mit einzubeziehen. Eine Theorie, die auf der Basis der Aktivierungstheorien den kognitiven Faktor – also die Funktion der Gehirnrinde – miteinbezieht, ist die Emotionstheorie von Schachter (ab 1960). Er konnte mit Hilfe von originellen Experimenten nachweisen, daß zwei Faktoren bei der Entstehung eines Gefühls von Bedeutung sind. Einerseits das bereits besprochene Aktivierungsniveau, das sich aufgrund von Außenreizen ändert. Dabei entsteht für den Organismus ein Zuordnungs- oder Benennungsbedürfnis. Nach Schachter determiniert der Aktivierungsgrad ausschließlich die Intensität einer Emotion. Die Qualität des Gefühls hängt von kognitiven Prozessen ab. Es kann also demnach der gleiche Aktivierungsgrad als Furcht oder Freude erlebt werden, je nachdem, wie die Umweltsituation interpretiert wird. Fehlt die physiologische Aktivierung oder ist sie sehr niedrig, wird es zu keiner Emotion kommen. Ein Beispiel ist die Beeinflussung des Erregungsniveaus durch Drogen. Ein Mensch, der unter sedierendem Drogeneinfluß in eine gefährliche Situation gerät, wird keine Furcht empfinden. Es entsteht andererseits ebenso keine Gefühlsreaktion, wenn die Aktivierung nicht auf kongnitiv-situative Faktoren, sondern auf somatische Ursachen zurückgeführt wird (z.B., wenn die erhöhte Aktivierung durch Einnahme von Medikamenten erklärt werden kann). Nur wenn keine derartige „direkte“ Erklärung für die gesteigerte physiologische Aktivität vorliegt, wird die Erregung in Zusammenhang mit der Gesamtsituation als Gefühl interpretiert.

Die Kritik an den Gedankengängen von Schachter bezieht sich auf den Hinweis, daß es emotionale Reaktionen gibt, die zu schnell ablaufen, um einen Bewertungsprozeß zu gestat-

ten. Diese Tatsache spricht jedoch nicht gegen die vorher dargestellten Emotionstheorien. Mann kann – in Verbindung der beiden Elemente – einen „dialektischen Prozeß" der Entwicklung eines Gefühlsrepertoires postulieren: Neue Situationen werden kognitiv bewertet, die physiologische Reaktion wird dadurch zu Gefühlen gefärbt und dann gespeichert. Entsteht in bestimmten Lebenssituationen öfter die gleiche Emotion, gerät dieses Gefühl unter „Stimuluskontrolle". Es kann durch die Reizsituation sofort – ohne vermittelnde Kognition – hervorgerufen werden. Diese gelernten „emotionalen Einstellungen" gehören zur Lern- und Entwicklungsgeschichte eines Individuums. Ein Gedankengang, der durch Beobachtungen verifizierbar ist. Kinder haben, je nach Entwicklungsstand, noch undifferenzierte Gefühle; die Kulturabhängigkeit von Gefühlen ist auf die unterschiedlichen Lern- und Interpretationsvorgänge abgeschlossener Sozietäten zurückzuführen.

Die Kognition als wesentliches Element der Emotion! Schließt man sich den linguistischen Deterministen oder der Definition von Bertrand Russell an, der sagt „Sprache ist nicht einfach dazu da, Gedanken auszudrücken, sondern Gedanken zu ermöglichen, die ohne sie gar nicht existieren können", so führt das zu der Behauptung: Es kann nur das differenziert gefühlt werden, was differenziert gedacht wurde.

Wittgenstein sagt, „die Sprache ist die logische Grenze der Welt" (*TLP* 5.6). Nichtsdestoweniger sagt er aber auch, daß es das Unsagbare, von dem geschwiegen werden muß, gibt (*TLP* 6.522). Gibt es also differenzierte Emotion jenseits der Sprache? Die logischen Positivisten sagen, daß der Grund, weswegen geschwiegen werden muß, einfach der ist, daß nichts da ist, worüber gesprochen werden kann (Wovon man nicht sprechen kann, darüber muß man schweigen).

Kann man ergänzen: Was man nicht benennen kann, kann man auch nicht fühlen?

LITERATUR

Berlyne, D. E., „The reward value of indifferent stimulation", in: Jack T. Tapp (Hrsg.), *Reinforcement and behavior* (New York 1969).

Magoun, Horace Winchell: *The waking brain*, 2nd ed. (Springfield, Ill. 1963).

Olds, J. & Milner, P., „Positive reinforcement produced by electrical stimulation of septual area and other regions of the brain" *Journal of comparative physiological psychology*, 47 (1958), S. 419–427.

Rohracher, Hubert, *Einführung in die Psychologie*, 9. unveränd. Aufl. (Wien 1965).

Schachter, Stanley & Singer, J. E.: „Cognitive, social and physiological determinants of emotional state" *Psychological review* 69 (1962), S. 379–399.

Vitouch, Peter, „Emotion", in: Kagelmann und Wenninger (Hrsg.), *Medienpsychologie/e. Handwörterbuch in Schlüsselbegriffen* (Wien 1982).

* * *

DAS LEIB-SEELE-PROBLEM AUS EVOLUTIONÄRER PERSPEKTIVE

Franz M. Wuketits
Universität Wien

Das Leib-Seele Problem (Materie-Geist-Problem) zählt zu den altehrwürdigen Problemen der Philosophie; nicht minder hat dieses Problem auch in anderen Disziplinen Kontroversen verursacht, so in der Psychologie, die ja der „Psyche" ihren Namen verdankt. Das Problem stellt sich meist in der Frage nach den möglichen Beziehungen zwischen Leib und Seele (Materie und Geist) dar. Nach Bunge (1980) kann man grundsätzlich drei Haltungen diesem Problem gegenüber einnehmen: Man kann es als Scheinproblem abtun; man kann es aber auch als echtes, jedoch unlösbares, oder schließlich als zwar echtes, allerdings lösbares Problem behandeln.

Wie ich an anderer Stelle (Wuketits, 1985) ausgeführt habe, betrachte ich das Problem keineswegs als Scheinproblem, halte es aber für prinzipiell lösbar – jedoch nur unter der Voraussetzung der „richtigen" Fragestellung. Meines Erachtens kann das Problem nur als „Gehirn-Bewußtsein-Problem" sinnvoll formuliert werden: Ausgehend davon, daß alles Seelische und Geistige im letzten nur auf der Basis des (menschlichen) Gehirns bzw. Zentralnervensystems zustande kommen kann, ist zu fragen, wie denn dieses Gehirn (die „oberste Instanz", das Steuerungszentrum des Zentralnervensystems) hervorbringt, was wir als seelische und geistige Phänomene bezeichnen. Vorweg gesagt: Diese Phänomene müssen spezifische Systemeigenschaften des Gehirns sein; und da das Gehirn, wie jedes biologische Organ, ein Evolutionsprodukt ist, ist es legitim und konsequent, auch alles Seelische, Geistige an die Evolution zu binden, zu fragen, wie in der Evolution des Lebendigen schließlich auch solche Phänomene entstehen konnten.

Man hat im Zusammenhang mit dem Leib-Seele-Problem häufig einen *Kategoriefehler* begangen. Ryle (1966) veranschaulichte das Wesen solcher Fehler anhand des folgenden sehr treffenden Beispiels:

> A foreigner visiting Oxford or Cambridge for the first time is shown a number of colleges, libraries, . . . scientific departments and administrative offices. He then asks ‚But where ist the University? I have seen where the members of the Colleges live, where the Registrar works, where the scientists experiment and the rest. But I have not yet seen the University . . . It has then to be explained to him that the University is not another collateral institution . . . The University is just the way in which all that he has already seen is organized (Ryle, 1966; S. 17f).

So wie in diesem Beispiel könnte nun jemand, dem man das Gehirn, seine einzelnen Bestandteile – Nervenzellen, Synapsen usw. – gezeigt hat, fragen: „Aber wo ist das Bewußtsein, der Geist? Ich habe nun die Elemente eines Gehirns gesehen, aber ‚den Geist' habe ich nicht gesehen." Selbstverständlich kann er den Geist nicht gesehen haben; weil der Geist ein spezifischer Ausdruck, eine spezifische Eigenschaft des Gehirns ist, oder besser gesagt eine Eigenschaft, die das System „Gehirn" aus dem spezifischen *Zusammenwirken* seiner Elemente (aus deren Organisation) aufweist. Es geht hier also grundsätzlich um zwei Dinge: daß ein System nicht mit seinen Eigenschaften verwechselt wird und daß die Eigenschaften eines Systems nicht vom System selbst getrennt werden.

Die Vertreter einer *strikt materialistischen* Position machen den Fehler, daß sie Geistiges mit Materiellem buchstäblich gleichsetzen und die Existenz des Geistigen leugnen (was einer komischen Note nicht entbehrt, wenn man bedenkt, daß ja auch solche Materialisten mitunter „geistvolle" Traktate mit dem Resultat eben einer Leugnung alles Geistigen geschrieben

haben); die Protagonisten *dualistischer* Auffassungen wiederum trennen den Geist vom Materiellen (vom Gehirn), um dann z.B. – wie Popper und Eccles (1977) – die Wechselwirkungen zwischen Gehirn und Bewußtsein (Geist) zu diskutieren, was ungefähr dem Versuch gleichkommt, „die Bewegung" von einem fahrenden Auto zu trennen, um sich dann darüber zu unterhalten, welche Beziehung denn zwischen einem Automobil und seiner Bewegung besteht. Eine solche Unterhaltung ist indes sinnlos, weil die Bewegung eine Eigenschaft des Autos ist. Ich meine also, daß das Leib-Seele-Problem häufig unter seiner falschen Formulierung leidet.

Hier wird das Leib-Seele-Problem formuliert als die Frage nach der Entstehung der höheren Bewußtseinsleistungen in der Evolution des Menschen. So haben wir die Möglichkeit, unter Rückbezug auf die Ergebnisse der empirischen Wissenschaften (Neurobiologie, Verhaltensforschung, Evolutionsbiologie), die Entwicklung jener Strukturen und Mechanismen zu rekonstruieren, die in letzter Instanz zum Auftreten des spezifisch menschlichen *selbstreflexiven Bewußtseins* (mit allen seinen Entäußerungen in Sprache, Kunst, Moralität) geführt haben. In seiner althergebrachten Form, in der Dualität von Leib und Seele, Materie und Geist, besteht das Problem unter diesen Prämissen nicht mehr.

Neuerdings ist es die *evolutionäre Erkenntnistheorie*, die alle geistigen Phänome (mit Selbstreflexivität, Erkennen und Denken) als Resultate von komplexen Prozessen der organischen Evolution ausweist (vgl. Lorenz, 1973; Lorenz und Wuketits, 1983; Riedl, 1980; Vollmer, 1980; Wuketits, 1984 a, b). Die Version des Leib-Seele-Problems in der evolutionären Erkenntnistheorie läuft auf die *Emergenz-These* hinaus, die beispielsweise auch von Bunge (1980) vertreten wird und die besagt:

> Geist als Gehirnfunktion ist eine (neu auftretende) Systemeigenschaft, zu der es keine Vorstufen zu geben braucht . . . Ein System kann Eigenschaften aufweisen, die keines seiner Untersysteme – auch nicht in „Vorstufen" – aufweist . . . Die Eigenschaften eines Systems können sich also wesentlich (qualitativ!) von denen seiner Teile unterscheiden (Vollmer, 1980; S. 22).

Es wird mithin zugestanden, daß der Geist irreduzibel bleibt, d.h. nicht „aufgelöst" werden kann in der Materie schlechthin; jedoch waren die *Vorbedingungen* für das Auftreten des Geistes in der Evolution materieller Art, gegeben also in der Evolution des Gehirns des Menschen. Im einzelnen heißt das:

1. Die Evolution kann, auch in der „subhumanen" Sphäre als ein „erkenntnisgewinnender" (informationsgewinnender) Vorgang beschrieben werden. Alle Lebewesen (Tiere) nehmen aus ihrer Umwelt mittels Sinnesorgane (oder, bei Einzellern, diesen äquivalenten Strukturen, etwa abgegrenzte Zellbezirke zur Wahrnehmung von Lichtreiz) Informationen aus ihrer Umwelt auf und vermögen in arterhaltend zweckmäßiger Weise darauf zu reagieren.
2. Damit bedeutet Leben *Lernen*, Evolution insgesamt ein „Versuch-und-Irrtum-Spiel"; ohne Aufnahme und Verarbeitung von Information aus ihrer jeweils spezifischen Umwelt würden die Lebewesen nicht überleben können. Man kann also sagen, daß jeder Organismus über einen „Weltbildapparat" verfügt (Lorenz, 1973), bestimmte Strukturen seiner Außenwelt wahrzunehmen und zu verrechnen in der Lage ist.
3. In der Evolution, die insgesamt eine Komplizierung und Differenzierung von Strukturen und Funktionen bedeutet, haben sich bestimmte Organe und Organsysteme ausgebildet, die die Aufgabe der Verrechnung der Außenwelt übernehmen: Sinnesorgane, Nervensysteme, Gehirne. Diese Organe (Organsysteme) sind immer komplizierter, differenzierter geworden, der „Horizont", das „Weltbild" der Lebewesen hat sich also in der Evolution, im Ganzen, beständig erweitert.
4. Alle, wenn man so sagen darf, „Erkenntnis"-Leistungen der Lebewesen im subhumanen Bereich spielen sich im wesentlichen *vorbewußt* ab. *Bewußte* Erkenntnis tritt erst ab einem sehr späten Evolutionsniveau auf, dem Niveau der Hominiden („Menschenartigen"), und auch dort kann man von Bewußtsein im engeren Sinne wohl erst auf der Stufe des *Homo*

sapiens sprechen; auf dieser Stufe standen bereits die berühmten Neandertaler, von denen Funde überliefert sind, die auf Kulthandlungen, auf Vorstellungen von einem „Jenseits" schließen lassen, was jedenfalls das Vermögen einer bewußten Reflexion der eigenen Existenz belegt.

5. Da während der ganzen Evolution der Hominiden (in einem Zeitraum von 4 bis 5 Jahrmillionen) eine sukzessive Vergrößerung der Gehirnkapazität feststellbar ist (von etwa 500 cm^3 bis etwa 1500 cm^3) und die uns vom Menschen der Gegenwart bekannten geistigen Leistungen erst mit einer verhältnismäßig hohen Gehirnkapazität auftreten, liegt die Annahme nahe, daß die Vergrößerung des Gehirns – nicht nur als Prozeß der Massenvermehrung, sondern vor allem auch als Ausbildung neuer Gehirnzentren (z.B. Sprachzentrum), also neuer „Qualitäten" – die anatomisch-physiologische Basis der Entstehung des Geistigen war.

Alles in allem gelangen wir also zu dem Schluß: Das Geistige ist ein Systemphänomen, eine Systemeigenschaft des Gehirns, die sich auszubilden vermochte auf der Basis entsprechender materieller Vorgänge.

Kybernetik und Systemtheorie liefern heute Modelle, die das Entstehen neuer Systemeigenschaften im allgemeinen begreifbar machen (zusammenfassend dargestellt von Wuketits, 1981). Schließen sich – zunächst unabhängig voneinander existierende – Elemente zu einem Ganzen zusammen, dann kann dieses Ganze sozusagen schlagartig neue Eigenschaften entwickeln, die auf der Ebene der Elemente (Systemteile) noch nicht vorhanden sind. Aus der ganz spezifischen Zusammenschaltung (Integration) der Elemente des Gehirns (Nerven- bzw. Gehirnzellen oder Neuronen) konnte dieses System Gehirn neue, vorher nie dagewesene Eigenschaften entwickeln – die man also vergeblich auf der Ebene der Bausteine sucht.

Wir müssen allerdings zugeben, daß wir die einzelnen Integrationsmuster des Gehirns heute noch nicht zur Genüge kennen. Das bleibt der künftigen Forschung vorbehalten; vielleicht aber wird ein „kleiner Rest" der Frage, wie denn die Elemente unseres Gehirns zusammenwirken, um letztlich Geistiges zu erzeugen, für immer unbeantwortbar bleiben. Auch werden evolutionäre, kybernetisch-systemtheoretische Modelle von Bewußtseinsvorgängen uns vermutlich kein *allgemeines* Schema liefern können für die *individuellen* Leistungen des menschlichen Gehirns, für die dem Menschen eigene *Personalität*. Das Gehirn des Menschen ist ein dermaßen komplexes System, daß darin offenbar „genug Platz bleibt" für individuelle, nur mit sich selbst identischen Leistungen, die in keine starren Schemata zu pressen sind.

Andererseits gewinnen wir damit, daß wir das Geistige im Hinblick auf seine Entstehung in der Evolution des Menschen betrachten, neue theoretische Möglichkeiten, jenes altersgraue Problem der Beziehung von Leib und Seele (Materie und Geist) zu lösen – oder bescheidener gesagt: uns einer Lösung zu nähern. Die Trennung von Leib und Seele jedenfalls ist kein geeigneter Ausgangspunkt für eine Lösung des Problems. Nehmen wir indessen die Evolution als Ausgangspunkt, die prinzipiell charakterisiert ist durch das Entstehen von immer neuen Systemen und Systemeigenschaften, dann haben wir dem Geistigen den Schein eines Wunders genommen – trotz (noch) bestehender „Restprobleme". Der Tradition verpflichtete Philosophen mögen von einer evolutionären Annäherung an das Leib-Seele-Problem, von der „biologischen Formulierung" des Problem enttäuscht sein. Dann aber bleibt ihnen nur, mit aller Klarheit gesagt, die endlose Repetition von Fragen, die Scheinfragen sind, und die Gewißheit, ein Problem so zu formulieren, daß es a priori unlösbar ist – was dem „Geschäft" der Philosophie nicht gerade zur Ehre gereicht.

LITERATUR

Bunge, M., *The Mind-Body Problem. A Psychobiological Approach* (Oxford-New York-Toronto 1980).

Lorenz, K., *Die Rückseite des Spiegels. Versuch einer Naturgeschichte menschlichen Erkennens* (München-Zürich 1973).

Lorenz, K. und Wuketits, F. M. (Hrsg.), *Die Evolution des Denkens (München-Zürich 1983).*

Popper, K. R. und Eccles, J. C., The Self and Its Brain. An Argument for Interactionism (Berlin-Heidelberg-London-New York 1977).

Riedl, R., *Biologie der Erkenntnis. Die stammesgeschichtlichen Grundlagen der Vernunft* (Berlin-Hamburg 1980).

Ryle, G., *The Concept of Mind* (Harmondsworth 1966).

Vollmer, G., „Evolutionäre Erkenntnistheorie und Leib-Seele-Problem", *Herrenalber Texte* 23 (1980).

Wuketits, F. M., „Kybernetik, Gehirn und Bewußtsein", *Umschau* 81 (1981).

Wuketits, F. M., (Hrsg.), *Concepts and Approaches in Evolutionary Epistemology. Towards an Evolutionary Theory of Knowledge* (Dordrecht–Boston–Lancaster 1984a).

Wuketits, F. M., *Evolution, Erkenntnis, Ethik* (Darmstadt 1984b).

Wuketits, F. M., *Zustand und Bewußtsein. Leben als biophilosophische Synthese* (Hamburg 1985).

* * *

ZUR VEREINBARKEIT VON KRIPKES THEORIE DER NOTWENDIGEN IDENTITÄT MIT DER LEIB-SEELE-IDENTITÄTSTHESE

Andreas Bartels
Universität Gießen

Die meisten Philosophen, die Saul Kripkes Argument gegen die Leib-Seele-Identitätsthese unter die Lupe nahmen, entdeckten darin die verunglückte Anwendung einer akzeptablen semantischen Theorie. Nach dieser Theorie etablieren alle wahren Identitätsaussagen, deren Terme starre Designatoren sind, notwendige Identitäten. Läßt man diesen semantischen Hintergrund unberührt, so muß man, um das Kripke-Argument zu entkräften, bei den Zusatzprämissen ansetzen, mit deren Hilfe Kripke mentale Begriffe in die Terminologie der starren Designatoren übersetzt: 1. Mentale Ausdrücke wie ‚mein Schmerz' sind starre Designatoren und 2. mentale Ereignisse sind unabhängig von Gehirnereignissen vorstellbar (Cartesische Intuition). Tatsächlich könnte ohne diese Prämissen Kripkes Theorie der notwendigen Identität nicht zu einer Aporie der Leib-Seele-Identitätsthese führen. Nur wenn die Ausdrücke ‚mein Schmerz x' und ‚mein Gehirnzustand y' starre Designatoren sind, zeichnet Kripkes Semantik eine mögliche Identität x=y als notwendige Identität aus. Und nur wenn ein Auftreten von x unabhängig von der Existenz von y denkbar ist, kann die notwendige Identität von x und y (und damit ihre Identität überhaupt) schließlich zu Fall gebracht werden. Hält man sich konsequent an Kripkes nichtepistemische Konzeption modaler Ausdrücke, so lassen sich nicht nur die Angriffe gegen die Zusatzprämissen zurückweisen; es läßt sich zeigen, daß die Grundidee seines Arguments gegen die Absicht Kripkes sogar als Stützungsinstanz einer starken Identitätsthese verwendet werden kann.

Aus welchem Grund ist Kripke davon überzeugt, daß der Ausdruck ‚mein Schmerz' ein starrer Bezeichnungsterm ist? Die Referenz des Terms ‚Schmerz' wird durch eine wesentliche Eigenschaft des Referenten bestimmt, denn, so Kripke, „Finden Sie es auch nur irgendwie plausibel, daß diese Empfindung selbst hätte existieren können, ohne eine Empfindung zu sein?"[1] Mein Schmerz kann nur existieren als diese besondere Empfindungsqualität, d.h. der Ausdruck ‚mein Schmerz' bezeichnet in jeder möglichen Welt eben jenen Eindruck, den er in der aktualen Welt bezeichnet. Ein naheliegender Einwand geht dahin, daß Schmerzen noch andere wesentliche Eigenschaften besitzen könnten; z.B. könnten sie, wie Blumenfeld[2] bemerkt hat, wesentlich mit bestimmten Gehirnzuständen identisch sein. Wenn die Referenz des Designators ‚mein Schmerz' nur durch *eine* wesentliche Eigenschaft des Referenten, nämlich durch seine Schmerzhaftigkeit bestimmt wird, so gebe es eine mögliche Welt, in der dem Referenten andere wesentliche Eigenschaften abgehen, z.B. die, ein Gehirnzustand zu sein. Der Referent von ‚mein Schmerz' wäre in unserer Welt wesentlich schmerzhaft und ein Gehirnzustand, in einer anderen Welt aber würde mit demselben Ausdruck eine andere Art schmerzhafter Ereignisse bezeichnet, die nicht Gehirnereignisse sind. Der Ausdruck ‚mein Schmerz' wäre folglich kein starrer Designator. „Mögliche Welten" fungieren in diesem Argument als Situationen, in denen Gegenstände auftreten, die zu einem Gegenstand der aktualen Welt ähnlich sind, in der Weise, daß sie einen oder mehrere wesentliche Aspekte des aktualen Gegenstandes teilen. Diese Sicht auf mögliche Welten ist aber nichts anderes als der von Kripke abgelehnte „telescope-view": wir richten im Geiste unser Fernrohr auf Realisationen von Beschreibungen und beurteilen ihre Ähnlichkeit zu aktualen Gegenständen. Für Kripke dagegen besitzen mögliche Welten keinen epistemischen Status (d.h. sie realisieren nicht Beschreibungen) und ihre Struktur ergibt sich nicht aus einer Ähnlichkeitsrelation zur aktualen Welt. Statt mit Beschreibungen ist der Begriff der möglichen Welten hier mit den Gegenständen selbst verankert und der Rahmen des Möglichen durch die Menge der gedanklichen

Inszenierungen bestimmt, die wir mit diesen Gegenständen veranstalten können, wenn wir uns dabei an die bekannten Naturgesetze halten. Deshalb vererben sich die wesentlichen Eigenschaften eines Gegenstandes in alle Kripke-möglichen Welten.

Eine zu Blumenfelds Kritik komplementäre Strategie verfolgt Levin,[3] indem er die Eignung von unmittelbaren phänomenologischen Qualitäten zur Referenzfixierung in Frage stellt. Seinen Angriff gegen den bei Kripke vermuteten phänomenologischen Essentialismus führt er mit dem Argument, daß innere Qualitäten nicht zur Erzeugung individuierender Beschreibungen verwendet werden können. Wenn wir mehr wollen als Schmerzen nur zu empfinden, also z.B. auf sie Bezug nehmen wollen, so müssen wir irgendetwas sagen, das auf sie zutrifft. Die einzigen Fakten aber, die uns über Schmerzen bekannt sind, betreffen die Umstände, unter denen sie auftreten. Folglich sind wir gezwungen, zur Referenzfixierung auf „äußere" Terme zurückzugreifen. Ebenso plausibel wie Levins Bemerkungen über die Bedingungen intersubjektiv zu rechtfertigender Bezugnahme auf mentale Ereignisse ist aber, daß wir auf unsere Schmerzen in einer Art Bezug nehmen können, die es uns erlaubt, sinnvoll von einer hypothetischen Situation zu sprechen, in der z.B. derselbe Kopfschmerz, der tatsächlich erst nach der Prüfung auftrat, uns schon während der Prüfung plagt. Diese Kompetenz scheint unabhängig davon zu sein, daß wir in der Lage sind, intersubjektiv aufschlußreiche Beschreibungen dieses Kopfschmerzes zu liefern, die anderen erlauben zu beurteilen, ob ich die Schmerzterme korrekt verwende, bzw. ob ich über Kriterien zur Identifizierung von Schmerzen verfüge. Für mich kann mein Schmerz durch seine spezifische Schmerzhaftigkeit individuiert werden, auch wenn ich nicht im Besitz von Kriterien bin, die festlegen, unter welchen Bedingungen ein Schmerz zu einem späteren Zeitpunkt derselbe Schmerz sein wird wie der gegenwärtige. Die starre Designation, also die Referenzfixierung für alle möglichen Welten im Sinne Kripkes, verlangt aber auch gar keine Identifizierung nach Kriterien. Sie beschreibt stattdessen die Konsequenzen in unserem Sprachverhalten, wenn wir etwas, z.B. Schmerzen, in der Art eines Einzeldings benennen. Um einen Schmerz im Sinne Kripkes starr zu bezeichnen, genügt bereits die Fähigkeit, sich diesen Schmerz zur Zeit t unter geänderten äußeren Umständen vorstellen zu können. Kripke benötigt also keine Inkorrigierbarkeits-These für mentale Erinnerungen oder mentale Selbstbeschreibungen. Die Rekonstruktion, die Kripke von unserem Gebrauch der Identität gibt, ist wesentlich nicht-epistemisch. Referenz und Identität fungieren als primitive Relationen, mit deren Hilfe ein Sprachgebrauch erläutert wird. Daraus ist klar, daß kein Rekurs auf individuierende phänomenologische Qualitäten erforderlich ist, um Kripkes Behandlung mentaler Ausdrücke als starre Designatoren zu rechtfertigen. Kripkes Verwendung von ‚Schmerz' als Terminus für Schmerzempfindungen schließt andererseits nicht aus, daß ‚Schmerz' in anderen Kontexten anders, z.B. als eine Bezeichnung verwendet werden kann, deren Referenz durch Beschreibungen in Termini des Verhaltens fixiert wird. Will man herausfinden, was Schmerzen „in Wirklichkeit" sind, so benötigt man äußere Identitätskriterien der Art „Schmerz x ist diejenige Entität, die das Schmerzverhalten y hervorruft". In diesem Fall wird x durch y kontingent herausgegriffen und x spielt die Rolle eines Platzhalters für mögliche wissenschaftliche Entdeckungen. Kripke aber verwendet diesen Ausdruck ausschließlich als Designator, der gestattet, meine subjektive Schmerzempfindung als Einzelding zu fixieren und unbeschadet seiner Identität in abweichende äußere Umstände gedanklich einzubetten. Kripkes These ist also gegen den Augenschein nicht antiphysikalistisch im Sinne einer behaupteten Analyseresistenz mentaler Phänomene; sie behauptet lediglich, daß eine bestimmte Verwendung mentaler Ausdrücke, nämlich jene, durch die subjektive Empfindungen fixiert werden, unabhängig von möglichen Entdekkungen über die physikalische Natur des Schmerzes ist.

Die „Cartesische Intuition" erhält in Kripkes semantischer Rekonstruktion entscheidende begriffliche Bedeutung: Es gibt eine mögliche Welt, in der C-Faser-Erregungen auftreten, ohne daß die Schmerzen existieren, von denen diese physikalischen Ereignisse in der aktualen Welt begleitet werden. Folglich kann die Identität von Gehirnzuständen mit Schmerzen keine notwendige sein; dies aber ist die einzige Form der Identität, die zugelassen ist, wenn ‚C-Faser-Erregung' und ‚Schmerz' starre Bezeichnungsausdrücke sind. Ein naheliegender

Einwand gegen eine argumentative Würdigung der Cartesischen Intuition ist, daß es Intuitionen dieser Art immer gibt, bevor man entdeckt, daß Phänomene, die zuvor getrennt vorstellbar, weil begrifflich unterscheidbar waren, tatsächlich identisch sind. Dieser Einwand wird jedoch gegenstandslos, wenn man sich klar macht, daß die Cartesische Intuition nichts mit mangelndem Wissen über mentale Ereignisse zu tun hat. Selbst wenn man bereits akzeptiert hat, daß ein schmerzhaftes Ereignis neurophysiologisch erschöpfend analysiert werden kann, bleibt die Cartesische Intuition in Kripkes Version lebensfähig, wie erstaunlicherweise durch einige Beispiele D. M. Armstrongs belegt wird, die dieser zur Stützung der kontingenten Identitätsthese gegeben hat.[4] Ich kann der Überzeugung sein, daß meine Hand nicht schmerzt, obgleich der physiologische Schmerz vorhanden ist, der unter Normalbedingungen als Schmerz in der Hand erfahren wird. In diesem Fall löst der tatsächlich vorhandene Schmerz vielleicht ein bestimmtes Schmerzverhalten aus (z.B. ein Greifen mit der anderen Hand nach der schmerzenden Stelle), ohne daß ein Bewußtsein des Schmerzes entsteht. Armstrong deutet diese Situation so, daß eine physikalische Kontrollfunktion meines ZNS blokkiert ist, die eben nichts anderes ist als mein Schmerzbewußtsein. Anstatt dieses Beispiel für die Cartesische Intuition auszunutzen, könnte man natürlich einfach die Kontrollfunktion zur physikalischen Seite der Identität Gehirnvorgang = Schmerz hinzufügen. Man kann aber die Cartesische Intuition noch radikaler formulieren: Der physiko-chemische Zustand (einschließlich der Funktion, die die Kontrollinstanz realisiert) könnte isoliert von Personen oder Gehirnen auftreten. Die Verknüpfung mentaler mit physikalischen Zuständen scheint zumindest vom kontingenten Umstand der Existenz von Personen oder Gehirnen abzuhängen, die diese physiko-chemischen Zustände als ihre mentalen Zustände besitzen, also z.B. Schmerz fühlen können. Daß mentale Zustände nichts anderes sind als physiologische Vorgänge in uns, bedeutet nicht, daß diese physiologischen Vorgänge, die uns als mentale Zustände präsent sind, nicht anders hätten existieren können als in der Form, in der sie uns gegeben sind.

Leplin[5] hat auf die Abhängigkeit hingewiesen, die zwischen der Schlüsselrolle der Cartesischen Intuition in Kripkes Argument und der Fixierung auf das Standardbeispiel für wissenschaftliche Identifikationen „Wärme = Molekularbewegung“ besteht. Ein Modell, bei dem wie im Fall Schmerz/Gehirnzustand sowohl der reduzierende wie der reduzierte Begriff durch wesentliche Eigenschaften herausgegriffen wird, ohne daß die Reduktion den Charakter einer kontingenten wissenschaftlichen Entdeckung verliert, ist die theoretische Realisation des Begriffs ‚Gen‘ durch das DNA-Molekül. DNA-Moleküle sind nicht eo ipso notwendig „Informationsträger“, sondern diese Eigenschaft kommt im kontingenten Zusammenhang eines Organismus mit gewissen Ablesefunktionen zustande. Auf dieselbe Weise, so Leplin, sind Gehirnzustände nicht notwendig Schmerzempfindungen, sondern erst aufgrund bestimmter qualifizierender Randbedingungen (einer bestimmten Art der neuronalen Verschaltung, einer speziellen Kontrollfunktion etc.), die nur kontingenterweise gegeben sind. Die starre Designation von ‚Schmerz‘ erweist sich dadurch als Artefakt, der durch die Reduktion verschwindet. Die vorher für wesentlich gehaltene Eigenschaft von Schmerzen, ihre spezifische Schmerzhaftigkeit, wird durch die Reduktion zur kontingenten Eigenschaft des reduzierenden physikalischen Zustandes.

Das Leib-Seele-Problem weist jedoch noch einen wesentlichen Unterschied zur Situation in Leplins Reduktionsmodell auf: Daß Schmerzen durch Empfindungsbegriffe essentiell herausgegriffen werden, wird nicht in gleicher Weise durch die Reduktion Schmerz/Gehirnzustand korrigiert, wie dies für die Eigenschaften des Vererbungsmechanismus gilt, die den Genen zugesprochen werden. Essentialität in diesem zweiten Fall wird gemessen an der Zentralität der Eigenschaft innerhalb der reduzierenden Theorie. Die Eigenschaft aber, bezüglich derer Schmerzen herausgegriffen werden, ist keine zentrale Eigenschaft einer Theorie (oder auch nur einer Theorienskizze) und zwar deshalb, weil es bei der Erforschung des Schmerzes nicht um sukzessive verbesserte Beschreibungen des Schmerzempfindens geht. Jeder Bezug auf die Vererbungsmechanismen erfordert eine grobe Beschreibung der Eigenschaften von Genen, während der Bezug auf Schmerzerlebnisse keine Beschreibungen impliziert. Damit läßt sich Kripkes Argument gegen die Identitätsthese aufrecht erhalten, soweit es um Schmerz

als Schmerzempfindung geht. Schmerzempfindungen stehen mit Gehirnzuständen deshalb nicht in der Relation der Identität, weil wir uns ihrer nicht mittels Beschreibungen vergegenwärtigen müssen, also nicht mittels individuierender Eigenschaften, die den Startpunkt für eine wissenschaftliche Reduktion abgeben könnten. Als identisch – und zwar als notwendig identisch – können sich dagegen ‚Schmerz' (in seiner Rolle als Erklärungsbegriff) und ‚Gehirnzustand' herausstellen. Auch ein vom Organismus isolierter Gehirnzustand realisiert die Rolle der Erklärung von Schmerzzuständen in Organismen, d.h. die Cartesische Intuition verliert hier ihre Wirksamkeit.

Daß die Anwendung der Kripke-Semantik auf mentale Begriffe eine notwendige Identität von Schmerz und Gehirnzustand etablieren kann, verschafft der Identitätstheorie einen entscheidenden Vorteil. Dem Dualisten wird der Ausweg versperrt, die Identität für die aktuale Welt zu akzeptieren, aber die Option nichtphysikalischer Schmerzen für mögliche andere Welten offenzuhalten. In dieser Sicht könnte Schmerz der Intension nach etwas wesentlich Unphysikalisches sein, das in unserer Welt zufällig mit physikalischen Sachverhalten extensional zusammenfällt. Die Kripke-Semantik dagegen verbietet es, das Wesen einer Sache gegen ihre empirische Realisation auszuspielen: Die empirische Realisation des Schmerzes in unserer Welt bestimmt das Wesen des Schmerzes in allen Welten, d.h. in allen mit unserem Wissen verträglichen Geschichten, die wir uns mit den Gegenständen unserer Welt ausdenken können.

ANMERKUNGEN

1 Kripke, S. A., *Name und Notwendigkeit* (Frankfurt 1981), S. 167.
2 Blumenfeld, J. B., „Kripke's Refutation of Materialism", *Australasian Journal of Philosophy* Bd. 53, Nr. 2 (1975), S. 151–156.
3 Levin, M. E., „Kripke's Argument against the Identity Thesis", *The Journal of Philosophy* Bd. 72, Nr. 6 (1975), S. 149–167.
4 Armstrong, D. M., *A Materialist Theory of the Mind* (London 1968).
5 Leplin, J., „Theoretical Identification and the Mind-Body Problem", *Philosophia* Bd. 8 (1979), S. 673–688.

* * *

VISION, BRAIN, AND MIND

Wolfgang Wenning
Berlin

Vision conveniently provides information on what is where in the world; objects as well as their positions or motions are identified without the effort of touching. Most higher animals derive their spatio-temporal orientation in the environment predominantly from vision. Bats are a well-known exception in managing to fly and hunt in the dark by listening to the echo of their own ultrasonic voices, similar to the technique of radar. One of the earliest attempts to explain vision is based on a similar principle. Empedokles (483–424) assumed that the eyes send out emanations which bring visual messages back to the eyes, as they return from external objects. Empedokles is known as well for the opposite view that effluvia evaporating from the pores of body surfaces stimulate vision when they reach the eyes. Demokritos, dispensing with emanations or effluvia, proposed that vision mirrors objects in the eyes. Aristotle rejected this view, stating: "It is strange too, that it never occured to him to ask why, if his theory be true, the eye alone sees, while none of the other things in which images are reflected, do so" (*On Sense and the Sensible*, Oxford, 1908). In his own theory Aristotle regards the eyes as an offshoot of the brain where vision occurs. A more detailed theory based on precise descriptions of the anatomy of the visual chain has been proposed by the early physiologist Galen of Pergamon (129–200). According to his theory, "animal spirits" (πνεύματα) originate in the brain, travel through the hollow tubes of the optic nerves, fill the crystaline lense, and from there, after interacting with infalling light, bring visual messages back to the brain on the same way . In short, to receive information about the world, the brain sends out messengers and listens to their reports when they return. This exchange of messages does not occur between the eyes and the world (as in the emanation theory) but between the brain and the eyes. Remarkably, Galen's model of "animal spirits" travelling back and forth between the brain and the eyes fits well to modern findings of feedback loops of afferent and efferent neural connections in the visual pathway. Indeed, the old model of information pick-up by visual spirits anticipates modern views on feature extraction from visual scenes.

In the Renaissance Kepler applied geometrical optics to the human eye and proposed the retina, rather than the lense to be the interface of visual information processing. Descartes corrected Kepler's optical analysis by using the sinus law of refraction and by attributing visual accomodation to changes in the shape rather than the position of the lense. In addition Descartes conceived of a geometrical mapping from a visual scene to the pinal gland of the brain, where the two inverted retinal images merge into an upright picture. Remarkably, Descartes was aware that this proposal shifts the problem of vision only a step deeper into the head, since it is not ". . . by means of this resemblance that we see the objects; for there is not another set of eyes inside the head with which we could see the images." (*La Dioptrique* (1637)). This argument is equivalent to Aristotle's argument against Demokritos who supposed that vision is mirroring. Indeed this argument seems to defeat any theory which explains vision by what Kepler, with respect to the retina, had called a "representative image". Descartes felt that the transport of retinal images through the optic nerves to the pinal gland of the brain does not solve the problem of vision without an additional step. To explain vision Descartes proposed the "substantia" of the *soul* to be located in the pinal gland such that vision occurs not between the eyes and the brain but rather between brain and mind. Descartes deviated from Galen's more correct anatomical description of the visual pathway with respect to the role of the nerve crossing at the optic chiasm as well as to what Galen called the thalamus

(known today as the thalamic lateral geniculate nucleus from where neurons project to the primary visual cortex); but he followed Galen in accepting the "animal spirits" (πνεύματα) to be the essence of the soul and in explaining the movement of the body by the conduction of these spirits through the hollow tubes of the nerves to the muscles which get, according to Descartes, inflated by this "very quick wind". Today the processes of neural conduction which cause motion and stimulates vision are understood to be based on excitation and inhibition. The problem remains how to bridge the gap of Cartesian dualism.

In his theory of vision Descartes exploited dualism in proposing that our ideas about the distance, size, form as well as position or motion of external objects are innately correct. That is to say the soul solves all the unsolved problems of vision. Not much remains to be explained if nativism is carried to extremes. In contrast to nativism, Hobbes and Locke revived the Stoic "tabula rasa" concept of the mind. Locke elucidated the empiricist view by a visual "Gedankenexperiment" which both Locke and Berkeley trace back to Molineaux: "Suppose a man born blind and now adult, and taught by his touch to distinguish between a cube and a spere [. . .]. Suppose then the cube and the sphere placed on a table, and the blind man be made to see; [. . . could he] now distinguish and tell which is the globe, which is the cube?" (Locke 1690). A Cartesian answer would be "yes"; Locke's answer—as well as Berkeley's —was "no"; the modern answer is a somewhat longer story.

To overcome dualism, Spinoza proposed: "ordo et connexio idearum idem est ac ordo et connexio rerum" (Ethica, 1677); Leibniz sophisticated this metaphysical approach in the Monadology; Berkeley, in his theory of vision, tried to overcome dualism by carrying the empiricist position to the extreme. He argued that distance cannot be immediately "seen", because the line connecting the distal stimulus at the object with the proximal stimulus at the retina is mapped on a point. It follows that, what appears to be a visual experience of distance (and similarly of size, form, position, and motion of external objects), must be a result of the process of thinking and learning, i.e. an activity of the mind. Further arguments concerning the other senses destroy the concept of space and time perception, and let the "real" world vanish in the mind. To explain what Kant later postulated to be the "synthetic unity of apperception" Berkeley proposed that the mind is part of God and that what appears as the real world does so only in God's perception.

In the 19th century, the metaphysical empiricism/nativism issue dissolved into research programs for a science of vision. Fechner distinguished between "outer psychophysics" which relates stimuli to "nerve energies" (J. Müller), and "inner psychophysics" which relates neural processes to sensations. A dualist in his epistemology, Fechner was a (panpsychic) monist metaphysically. Mach was also a monist, while epistemologically he believed in psychophysical parallelism. In "Über die Wirkung der räumlichen Verteilung des Lichtreizes über die Netzhaut" (1865) Mach states a principle of psychophysical parallelism which he regards to be "ein heuristisches Prinzip der psychophysischen Forschung". Hering underlines this in "Zur Lehre vom Lichtsinne" (1878) by calling Mach's principle of psychophysical parallelism a "conditio sine qua non aller solchen Forschung, wenn sie Früchte tragen soll". Modern research programs in the fields of psychophysics, neuro-physiology, and phenomenology of vision are foreshadowed by this view.

Let me switch to neurobiology and ask if the mind may be caught within the network of neurons in the brain. Let me approach this question by relating theories on (1) *color*, (2) *distance*, and (3) *form* perception of the 19th century to recent results in neurobiology.

(1) Thomas Young in 1806 conjectured that three different light sensitive "particles" or "resonators" in the human retina would account for the outcome of Newton's additive color mixture experiments. Helmholtz adopted and further developed this idea by proposing absorption curves for three kinds of hypothetical photoreceptors which give rise to a three-dimensional color space. The Young-Helmholtz tristimulus theory of color vision found wide acceptance in psychophysics long before Wald in 1965 photochemically confirmed Young's original hypothesis by identifying three types of photo-pigments in retinal cones (and a fourth one in the rods for scotopic vision). It is obvious that Young applied some principle of psycho-

physical correspondence in his conclusion that lead from the psychology of additive color mixture to the physiology of the retinal surface.

Ewald Hering in 1878 proposed a different model of color vision which turned out to be correct at a deeper level of neural processing. Hering assumed that there are three opponent color systems of antagonistic physiological processes: the black/white, the red/green, and the yellow/blue-system. In each system, Hering supposed, the antagonistic mechanism consists of two kinds of activities which cancel each other at an equilibrium point. Physiologists predominantly rejected Hering's theory until around 1960, when electrophysiological techniques of single-cell recordings led to the discovery of opponent color systems in the visual pathway. Hubel and Wiesel showed that at the level of the primary visual cortex so-called simple neurons encode bright versus dark, red versus green and yellow versus blue antagonistically by increasing/decreasing firing rates depending on excitatory/inhibitory influences of concentric ON/OFF receptive field areas. (Red-ON, Green-OFF)-cells are excited by red and inhibited by green light. (Yellow-ON, Blue-OFF)-cells are excited by yellow and inhibited by blue light. Recently Michael 1978 identified color contrast neurons which get maximally stimulated if one opponent color hits the center and the other hits the suround of the dual opponent color receptive fields of these neurons. It is noteworthy that color contrast phenomena had originally motivated Hering's theory.

Let me mention the phenomenon of color constancy which puzzled theorists of vision especially since Katz 1911. Helson and Judd in the thirties found a rather complicated psychophysical description for it. In the sixties Land explained color constancy by his retinex theory of color vision elaborated later by Horn. Land, Hubel et al. 1983 conclude from split-brain experiments that the Horn-algorithm is performed behind the retinal level. Kolata 1982 reports the recent discovery that certain neurons in the primary visual cortex "do exactly what is needed for the sort of color perception that Land described".

(2) Since Thomas Reid 1785 theorists of vision in the empiricist tradition have distinguished sharply between "sensation" (e.g. of colors) and "perception" (e.g. of distance). Helmholtz, similar to Berkeley, based his theory of distance perception on the concept of "unconscious judgements and conclusions" performed by the mind. On the other hand, Hering in his nativist approach to distance perception proposed that color sensation and distance perception are both immediate experiences based on physiological mechanisms of similar complexity. Today the old problems raised in the empiricism/nativism debate between Helmholtz and Hering begin to find answers, such as the question of whether object recognition precedes or is preceded by distance perception: Bela Julesz, using his technique of random dot stereogramms, showed that form perception may even be constituted by mechanisms of distance perception, although usually we use object recognition as a cue to estimate distance. Visual science today buries the empiricism/nativism issue in the technicalities of a modular description of visual skills (Marr 1982, Poggio 1984). A kind of "Molineaux experiment"—proposed by Locke and Berkeley to settle the issue—has been performed by Blakemore and Cooper 1970 with cats. If kittens are raised in vertically/horizontally striped cylinders, their later behaviour as well as their neural connectivities in the visual cortex will depend on the environment, although not completely.

(3) Theories of visual form perception emerged with the concept of "Gestalt" proposed by Ehrenfels 1890 under the influence of Mach and Brentano. In what later became known as the Gestalt-school of Wertheimer, Köhler and Koffka, Mach's principle of psychophysical parallelism reappeared in the form of the much stronger principle of "psychophysical isomorphism", according to which (quoting from Köhler 1929 in condensed form) experienced order in space and time is always structurally identical with a functional order in the sequence of distributions of correlated brain events, such that units of experience go with functional units of underlying physiological processes (p. 61–63). Köhler assumed that the varying patterns of retinal stimulation constrain and alter processes of dynamic field organization in the brain where the interaction of (electromagnetic) field forces (arising from ionic reactions) leads somehow to the formation of spatio-temporal configurations, that is, "Gestalten" which are

"isomorphic" with what is perceived. According to this somewhat "neodemocritean" model, vision is a kind of distorted, three-dimensional "mirroring" of visual forms in the brain. Köhler 1929 explicitely rejected the idea that visual form is represented in the brain "symbolically" by "local events" and criticized Bühler for holding this view:

> "Once K. Bühler tried to give an explanation of a very characteristic shape, that of the straight line. He assumed that all retinal points which form a straight line are anatomically connected in a special manner, and that this gives a straight line its particular appearance. This hypothesis has the character of a machine theory. I do not think that we can hope to solve our problem in this fashion. There are a great many highly characteristic shapes besides the straight line. Are we to assume that for each there is a special anatomical arrangement? Or rather, a great many such arrangements for each single shape, since each may be projected upon many different parts of the retina"?

Bühler's view that the retinal points forming a straight line are anatomically connected coincides exactly with the findings of Hubel and Wiesel. Along the visual pathway neurons widen their scope by surveying neural activity at earlier levels. Simple, complex and hyper-complex neurons of the primary visual cortex integrate and modulate incoming signals in columnar circuits to build up opponent color tuned feature detectors for dots, oriented lines, edges, rectangles and (abstracting from retinal position) movement of these shapes. Bühler theoretically anticipated these experimental results. On the other side, Köhler anticipated a modern theoretical problem which amounts to the question if there are "grandmother detection neurons" or "yellow VW detectors" (Dodwell) in the brain. Köhler's principle of psychophysical isomorphism reduces the complexity of this problem in the most simple way by continuity conditions connecting spatiotemporal units of experience with spatiotemporal units of underlying brain processes. As discovered by Hubel and Wiesel there is indeed a topological correspondence between patterns of retinal stimulation and patterns of neural activity in the primary visual cortex. The most advanced theory of vision today—Marr 1982—builds its basic concept of the "primal sketch" on this correspondence, reviving Kepler's idea of a "representative image" within the computational paradigm of vision. Descartes' objection (see above) had been that the idea of a "representative image" (or a homunculus in the head) reproduces the problem that it is supposed to solve. However, if there is more than one homunculus in the brain, each one may solve a smaller problem. If there is a hierarchical population of homunculi in the brain, visual competence may be performed at different levels of visual skills. Let me quote a passage from Trevarthen in this connection (see "Brain and Mind", CIBA Foundation, a.a.O.):

> "I am an isomorphist. I think there is a little man in the head. In fact I think there is a whole crowd of people. They all understand each other perfectly well, they are innate, they make decisions in a properly formed hierarchy, they disagree occasionally but they know how to resolve their disagreements, and they also have specialized occupations which are at least as sophisticated as anything that has been created by technology. I could show you a picture of a little 'man' in the brain of a fish, but all I want to say now is that the task of getting about in space is solved morphologically by all active animals, and even by active plants."

A computational model along these lines had been developed by Selfridge 1959 for visual pattern recognition. It is called "pandemonium" and consists of a hierarchy of "image demons", "computational demons", "cognitive demons" and "decision demons". Each "demon" performs a definite computational task. Together the "demons" extract more and more complex features of a visual scene. It may be said that each "lower" demon extracts a bit of geometric information and transforms it into an algebraic one, which is delivered to the "higher" demons until a "decision" demon settles part of a scene analysis by giving a binary reply (yes or no).

The basic idea of this approach fits into the more general context of connections between geometry and algebra. *Counting* (attributing a number to a set of objects) and *measuring* (attributing a number to e.g. a segment of a line) are prominent examples of such connections. In attributing coordinates (ordered pairs of numbers) to each point of a plane figure (point, line, circle, etc.) Descartes isomorphically represented geometry within algebra. A somewhat

more abstract connection of this kind has been discovered by Leonard Euler. He counted "vertices" (points), "edges" (lines) and "faces" of closed polyhedral surfaces, combined these numbers into a simple arithmetical expression (the so-called "Eulercharakteristik" χ = number of vertices minus number of edges plus number of faces), and in this way showed how to give a complete classification (up to topological isomorphism) of closed surfaces in three-dimensional space (e.g. the sphere, the torus, the double-torus, etc.) and also of non-orientable surfaces (e.g. Klein's bottle in four-dimensional space). Nowadays algebraic topology broadens and generalizes Euler's approach.

It seems that the brain is a pandemonium of mathematicians: more specifically, let the "image demons" be geometers concerned with geometric projections and transformations preserving certain types of invariants (stated in Klein's "Erlanger Programm"). "Computational demons" then will turn out to be algebraic topologists associating algebraic structures (numbers, groups, moduls etc.) and their interrelations (isomorphisms, homomorphisms) to the geometric output of the "image demons". In this way points, lines, circles, spheres, tori, etc. are related to algebraic entities. At the stage of "cognitive demons" we would encounter the species of pure algebraists. These perform calculations on algebraic structures which may be overwhelmingly complicated in their combinations but intellectually bound to only a few rules. At the top, finally, the "decision demons" apply even simpler rules concluding from algebraic data simply "true" or "false". These are the logicians who sometimes step out of their own personality and become metamathematicians to avoid (semantical) contradictions arising from problems of (conscious) self-reference.

Returning from the "metaphorical brain" (Arbib) of artificial intelligence to the ground of neurobiology, we do find connections. The first geometer in the visual chain is, of course, the lense of the eye which maps scenes onto the retina. The retinal output is an intrinsic combination of geometric and algebraic information: geometric information is reproduced topologically in the primary visual cortex, as evidenced by Hubel and Wiesel from measurements of cortical activity. Algebraic information is extracted from retinal signals in the visual cortex by local feature detectors for spots, oriented lines, angles, and movement of simple shapes (see above). Thus the brain intrinsically combines a geometric and an algebraic mode of information processing. It may well turn out that evolution found a clever way to combine both modes up to the level of visual decision making in recognition, and perhaps even up to the level of sentence creation out of mental images.

Visual science proceeds from phenomenological descriptions of visual abilities over psychophysical laws and models of functioning to neurobiological discoveries. Hering's theory of color vision is a paradigm case of this progression and has encouraged similar studies in distance and form perception. Cognitive abilities like attention, awareness, and consciousness however are not yet understood very well. According to Brentano and Husserl, consciousness is "directedness to something" or "intentionality". Creutzfeld 1979 suggests that if consciousness can be understood only in dualistic terms, it may be asked, whether "intentionality" has a neurophysiological basis. He distinguishes between "external loops" (thalamo-cortical circuits e.g. of visual information take up) and "internal loops" (feeding midbrain/cerebellum information into the neocortical areas) and connects both by "reflective loops" to describe the ability of the brain to "represent to itself its self-representation". Berkeley compared the brain with a swarm of bees. The challenge then is to describe the swarm in terms of mathematics and to describe mathematics in terms of the bees.

REFERENCES

Creutzfeld, O. D., "Neurophysiological Mechanisms of Consciousness", in: *Brain and Mind*, CIBA Foundation–Excerpta Medica (Amsterdam-Oxford-New York 1979).

Hering, E., *Zur Lehre vom Lichtsinne. Sechs Mitteilungen an die Kaiserl. Akad. der Wissenschaften.* (Vienna 1878).

Hochberg, J. E., "Nativism and Empirism in Perception", in: *Psychology in the Making*, L. Postman, ed. (New York 1962).

Köhler, W., *Gestaltpsychology* (New York 1929).

Marr, D., *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information* (San Francisco 1982).

Wenning, W., "Colors and Languages", in *Logics and Linguistics, Problems and Solutions. Selected Papers of the 7th Int. Congr. of Logic, Methodology and Philosophy of Science* (New York 1984).

* * *

3. Das Ich und die Welt

3. The Self and the World

FOUR HYPOTHESES ON THE ORIGIN OF MIND

Julian Jaynes
Princeton University

What I have to present to you this morning is essentially a historical theory that results in a conclusion that consciousness and mind cannot be understood apart from their history. To me, all discussions of mind must be diachronic, not synchronic as most such discussions are. It is a position which I recognize is not popular or even consonant with modern philosophical thinking, but is one which I would like to urge upon you whether or not you accept the details of what I am going to present.

The problem that I start with is that of the origin of consciousness in an evolutionary sense. It is a problem that after the rise of evolutionary theory in the latter part of the 19th century was then regarded as massively important, indeed, imperative for the completion of the theory of the origin of species by natural selection and thereby its vindication. Evolution explained the materialistic and chance origin and proliferation of species, but how could it explain the origin of the seemingly inmaterial human mind? That was and still is the problem. How could consciousness have arisen out of mere matter by natural selection?

Some, like Wallace, said it couldn't: consciousness in humans had to be imposed by a Deity. So did Darwin, surprisingly, but at the beginning of evolution as one of the several powers which were "originally breathed by the Creator into a few forms or into one", my quotation being from the last paragraph of the *Origin of Species*. And this Darwinian idea that conscious mind evolved in parallel with morphology led to a program of research into the evolutionary development of consciousness by Darwin's disciple, Romanes—a program which was quickly seen to be inadequate. Other attempted solutions followed, such as Huxley's implausible brand of adversarial proto-behaviorism or the mysticism and imprecision of emergent evolution, and then in philosophy with the neo-realism of Whitehead, Alexander, Perry, and others which is unfortunately having a modern revival as a result of some of the paradoxes of "observer-participancy" in quantum physics. And of course along this tortuous way, the psychologist's flippant solution to the problem of the origin of consciousness by denying that it exists at all. Coming through such an exhausting and trying history, it is not surprising that the problem has withered away its 19th century urgency and disappeared from both scientific and philosophical discussions. I think it should be brought back to clarify many of the problems being addressed at this symposium.

The reason for this abysmal failure to solve the problem of the origin of consciousness is because of a false and unwieldy notion of consciousness. Through history it had accumulated a huge amount of cultural baggage which would not allow the matter to be looked at properly. Some of this excess of reference was due to religious history wherein the soul became invested with all possible psychological functions which then became the properties of mind or consciousness as those words replaced the term soul. Another source of such error was due to philosophical thinking itself, that what we could arrive at as a function of mind by logical inference, that too was a part of consciousness even though such propositions had no introspective validity. So the first and proper step in this analysis is to try to loosen the concept of consciousness from the accretion of these historical mistakes.

By the infusion of logically inferred capacities into the idea of consciousness, I mean processes such as sense-perception. While we can be conscious, (though not always) of the products of senseperception, the actual processes are in no way accessible to conscious introspection. Historically, we inferred and abstracted these processes from a realization of our sense

organs, and then, because of prior assumptions about mind and matter or soul and body, we believed these processes to be a part of consciousness—which they are not. If any of you still think that sense-perception is to be equated to consciousness, then I think you could follow a path to a *reductio ad absurdum*: you would then have to say that since all animals have sense-perception, all are conscious, and so on back through the evolutionary scale even to one-celled protozoa, and thence to even the amoeboid white cells of the blood since they sense bacteria and devour them. They too would be conscious. And to say that there are ten thousand conscious beings per cubic milimeter of blood in each of us here this morning is a position few would wish to defend.

Now to some of you it will seem that I am defining consciousness in a much narrower way than is usual. I am. But it is part of the theory I am stating that that is what consciousness is, and not sense perceptions—as I have just stated. I do not mean that all consciousness is introspection, but it is introspectable. Consciousness is what is going on "in the minds" of any dozen people now on the street in Kirchberg, and if you should ask them to say into a tape recorder everything they are thinking about, you will not find bundles of sense-perceptions, but a stream of worries, regrets, hopes, reminiscences, interior dialogues, and monologues, plans, imaginations—all of which are the stuff of consciousness.

Why should it be that some people confound consciousness and sense perception? The reasons are several. First, as I shall be describing, consciousness is an anlalog of perception and so indeed can *seem* to be identical. I shall try to show that as the physical sense-organed I is to a physical object, so on the basis of language an analog 'I' is to an 'object' of consciousness. And a perception can indeed be an 'object' of consciousness though not always.

But it is surprising that sophisticated philosophers and psychologists, also confound sense perception and consciousness. In psychology, psychophysics began with this error. Fechner, a panpsychist, felt he was relating the entire universe of mind with that of matter by measuring the just noticeable differences in an attribute of sensation, and so forming the famous Weber-Fechner Law. Some psychologists even today who are studying perception think they are studying consciousness. As William James pointed out long ago, sense-perceptions are abstractions we impose on experience, not intrinsic parts of that experience itself. *His* starting point was only that 'thought goes on'. And so is mine.

Some modern philosophers also perpetuate this error. I imagine it is because of the old traditions about sense data. When Russell, looking for an example of consciousness, simply says, "I see a table", that is a highly artificial and misleading choice—like saying that a b flat is an example of a symphony. Yes, he is conscious of seeing a table (which incidentally he could probably do better with his eyes closed), but the seeing of the table is not consciousness but vision and an object of consciousness, even in the midst of what he was really conscious of, his argument. Russell should have picked a more ethologically valid example that was really true of his consciousness, that had really happened, such as "When will I find time to rework the *Principia*", or "How can I afford the alimony for the second Lady Russell". He would then have come to other conclusions. Such examples are consciousness in action. "I see a table" is not. Descartes would never have said, "I see a table, therefore I am."

Just as sense-perception is not due to consciousness, so with all the variety of preceptual constancies, as size, brightness, color, shape, which our nervous systems preserve under widely varying environmental changes of light, distance, angle of regard, or even our own moving about as objects retain their same position, called location constancy—all done without any help from consciousness.

So with another large class of activities we call *preoptive*, such as how we sit, walk, move. All these are done without consciousness, unless we decide to be conscious of them—the preoptive nature of consciousness. Even in speaking, the role of consciousness is more *interpolative* than any constant companion to my words. I am not now entering my lexical storehouse and selecting items to string on these syntatic structures. Instead, I have what can best be described as an intention of certain meanings, what I call *structions*, and then linguistic habit patterns take over without further input from my consciousness. Similarly, in hearing someone

speak, what are listeners conscious of? If the flow of phonemes or even the next level up of morphemes or even words, you would not be understanding what I am intending.

That consciousness is in everything we do is an illusion. Like asking a flashlight to turn itself on and look around a completely dark room to see if there is light, the flashlight would conclude that the room was everywhere bright with light—which would be false.

A further error about consciousness stems from the beginning of empiricism when Gassendi used the Aristotelean term *tabula rasa*, and which Locke, who was basing the first two books of his *Essay* on Gassendi's *Syntagma Philosophicum*, anglicized as the mind's "white paper, void of all characters, without any ideas."(*Essay* II, 1.2). Had the camera been around at the time, I suggest Locke would have used it instead as his foundational metaphor. In experience, we take successive pictures of the world, develop them in reflection, bringing into existence concepts, memories, and all our mental furnishings.

But that consciousness does not copy experience can be shown very easily: (a) by examining the memories we have and noting that they are not structured the way we experienced them, e.g. thinking of the last time you were in swimming; most people, instead of the complicated multi-sensory experience as it actually was, tend to see themselves swimming from another point of view—something of course they have never experienced at all; or (b) by examining the absence of memories that we should have if consciousness did copy experience, e.g. knowing what letters go with what numbers on telephones which we have stared at thousands of times, or to sketch from memory while here in Kirchberg the main entrance-way to your office building—you will find it difficult. Conscious memory does not copy experience but reconstructs it as a must-have-been.

Empiricism is also the culprit in the belief that consciousness is necessary for learning. In the old terminology, mind had free ideas, copies of what was perceived, which when things were perceived together, their copies or ideas stuck together—the association of ideas. And since this presumably went on in consciousness, it seemed clear that learning or the association of ideas was the crucial criterion of consciousness. This was explicitly stated by many early psychologists. But a huge body of experimental research in later psychology suggests that this is far from true. Learning can be found in species and preparations where no one would think it plausible that something like introspective consciousness has anything to do with it whatever. In humans, all types of learning, conditioning, motor learning, and instrumental learning or operant conditioning can be shown to occur without any awareness or assistance from consciousness. This is not to say that consciousness does not play a role in human learning, as in decisions as to what to learn, or making rules of how to learn better, or consciously verbalizing aspects of a task. But this is not the learning itself. And my point is that consciousness is not necessary for learning to occur.

One could here bring up the well-known phenomenon of the automatization of habit, for when this happens to us, it seems that the task has required consciousness at the beginning, but as the habit is perfected consciousness eases away and the task is performed effortlessly. This same smoothing out and increased rapidity of a habit with practice is universal among all animals that can learn. Generally, in this ubiquitious phenomenon, it is not the lapsing of consciousness with improved performance so much as the lapsing of forced attention to components of the task. And attention, which is the focusing of sense-perception, is not necessarily conscious. If you take two coins in either hand, and toss them across each other so that the opposite hand catches each, this is a task that will take somewhere between 15 and 20 trials to learn. And if you wish to try this this evening, and monitor your consciousness while you are doing so, you will find that consciousness has little to do with the learning that seems to go on mechanically. You might be conscious of something about your clumsiness, or the silliness of what you are doing as you keep picking the coins up from the floor, until, at the point of success, your consciousness is somewhat surprised and even proud of your superior dexterity. It is the attention which has changed. Automatization is a diminution of attention not consciousness.

Consciousness is also not necessary for thinking or reasoning, a somewhat shocking state-

ment. But in 1901, Karl Marbe as a graduate student at Würzburg performed the most simple and yet most profound experiment of his day. Using his professors as subjects, each of whom was highly experienced in introspective experiments, he asked them to make a simple judgment between two identical looking weights as to which was the heavier. Against the background of the experimental psychology of the time, the result was astonishing. There was no conscious content for the actual judgment itself, although such a judgment was embedded in the consciousness of the problem, its materials, and technique. So began what came to be called the Würzburg School of Imageless Thought, which lead through experiments by Ach, Watt, Kulpe and others to concepts such as set, *aufgabe*, and determining tendency—which I have renamed structions. Structions are structures like instructions given to the nervous system, that, when presented with the materials to work on, result in the answer automatically without any conscious thinking or reasoning. This phenomenon applies to most of our activities, all the way from such simplicities as judging weights to solving problems to scientific and philosophical activity. Consciousness studies a problem and prepares it as a struction, a process which may result in a sudden appearance of the solution as if out of nowhere. During World War II, British physicists used to say that they no longer made their discoveries in the laboratory; they had their three B's where their discoveries were made, the bath, the bed, and the bus. And this process on a smaller scale is going on in me at present as my words are as if chosen for me by my nervous system after giving it the struction of my intended meaning.

Finally in this list of mistakes about consciousness, a word about its location. Most people, with possibly the present company excepted who have thought long about the problem, and so placed it 'out there' in the intellectual domain, tend to think of their consciousness, much as Descartes, Locke and Hume did, as a space usually located inside their heads. There is of course no such actual space whatever. The space of consciousness, which I shall hereafter call mind-space, is a functional space that has no location except as we assign one to it. To think of our consciousness as inside our heads is, because of all our words like introspection, internalization, etc. a very natural but arbitrary thing to do. I certainly don't mean to say that consciousness is separate from the brain; by the assumptions of natural science, it is not. But we use our brains in riding bicycles, and yet no one considers that the location of bicycle riding is inside our heads. The phenomenal location of consciousness is arbitrary.

To sum up so far, we have shown that consciousness is not all mentality, not to be equated with sensation or perception, that it is not a copy of experience, nor necessary for learning, nor even necessary for thinking and reasoning, and has only an arbitrary and functional location. As a prelude to what I am to say later, I wish you to consider that there could have been at one time human beings who did most of the things we do, speak, understand, perceive, solve problems, etc. but who were without consciousness. I think this is a very important possibility.

So far this is almost going back to a radical behaviorist position. But what then is consciousness, since I regard it as an irreducible fact that my introspections are as real as so-called "external" perceptions, though with a distinguishably incongruent quality? My procedure here will be to outline in a somewhat terse fashion a theory of consciousness and then to explain it in various ways.

Subjective conscious mind is an analog of what we call the real world. It is built up with a vocabulary or lexical field whose terms are all metaphors or analogs of behavior in the physical world. Its reality is of the same order as mathematics. It allows us to short-cut behavioral process and arrive at more adequate decisions. Like mathematics, it is an operator rather than a thing or a repository. And it is intimately bound with volition and decision.

Consider the language we use to describe conscious processes. The most prominent group of words used to describe mental events are visual. We 'see' solutions to problems, the best which may be 'brillant' or 'clear' or possibly 'dull', 'fuzzy', 'obscure'. These words are all metaphors, and the mind-space to which they apply is generated by metaphors of actual space. In that space we can 'approach' a problem, perhaps from some 'view-point' and 'grapple' with its difficulties. Every word we use to refer to mental events is a metaphor or analog of some-

thing in the behavioral world. And the adjectives that we use to describe physical behavior in real space are analogically taken over to describe mental behavior in mind-space. We speak of the conscious mind as being 'quick' or 'slow', or of somebody 'nimble-witted' or 'strong-minded' or 'weak-minded' or 'broad-minded' or 'deep' or 'open' or 'narrow-minded'. And so like a real space, something can be at the 'back' of our mind, or in the 'inner-recesses' or 'beyond' our minds.

But, you will remind me, metaphor is a mere comparison and cannot make new entities like consciousness. A proper analysis of metaphor shows that it can. In every metaphor there are at least two terms, the thing we are trying to express in words, the *metaphrand*, and the term produced by a struction to do so, the *metaphier*. These are similar to what I. A. Richards called the tenor and the vehicle, terms more suitable to poetry than to philosophical analysis. I have chosen metaphrand and metaphier instead to have more of the connotation of an operator by echoing the arithmetic terms of multiplicand and multiplier. If I say the ship plows the sea, the metaphrand is the way the bow goes through the water and the metaphier is a plow.

As a more relevant example, suppose we are a person back at the formation of our mental vocabulary, and we have been trying to solve some problem or to learn how to perform some task. To express our 'See' is the metaphier, drawn from the physical behavior from the physical world, that is applied to this otherwise inexpressable mental occurrence, the metaphrand. But metaphiers usually have associations that we are calling *paraphiers* that then project back into the metaphrand as what are called *paraphrands*, and, indeed, creating new entities. The word 'see' has associations of seeing in the physical world and therefore of space, and this space then becomes a paraphrand as it is united with this inferred mental event called the metaphrand.

metaphrand	→	metaphier
‖		↓
paraphrand	←	paraphier

In this way the spatial quality of the world around us is being driven into the psychological fact of solving a problem (which as we remember needs no consciousness). And it is this associated spatial quality that, as a result of the language we use to describe such psychological events, becomes with constant repetitions this functional space of our consciousness, or mind-space. This mind-space I regard as the primary feature of consciousness. It is the space which you preoptively are 'introspecting on' or 'seeing' at this very moment.

But who does the 'seeing'? Who does the introspecting? Here we introduce analogy, which differs from metaphor in that the similarity is between relationship rather than between things or actions. As the body with its sense organs (referred to as I) is to physical seeing, so there develops automatically an analog 'I' to relate to this mental kind of 'seeing' in mind-space. The analog 'I' is the second most important feature of consciousness. It is not to be confused with the self, which is an object of consciousness in later development. The analog 'I' is contentless, related I think to Kant's transcendental ego. As the bodily I can move about in its environment looking at this or that, so the analog 'I' learns to 'move about' in mind-space concentrating on one thing or another.

A third feature of consciousness we are calling *narratization*, the analogic simulation of actual behavior, an obvious aspect of consciousness which seems to have escaped previous synchronic discussions of consciousness. Consciousness is constantly fitting things into a story, putting a before and an after around any event. This feature is an analog of our physical selves moving about through a physical world with its spatial successiveness which becomes the successiveness of time in mind-space. And this results in the conscious conception of time which is a spatialized time in which we locate events and indeed our lives. It is impossible to be conscious of time in any other way than as a space.

There are other features of consciousness which I shall simply mention: *concentration*, the 'inner' analog of perceptual attention; and its opposite, *suppression*, by which we stop being conscious of annoying thoughts, the analog of turning away from annoyances in the physical

world; *excerption*, the analog of how we sense only one aspect of thing at a time; and *consilience*, the analog of perceptual assimilation; and others. In no way is my list meant to be exhaustive. The essential rule here is that no operation goes on in consciousness that was not in behavior first.

Cognitive psychologists are sometimes justly accused of the habit of reinventing the wheel and making it square, and then calling it a first approximation. I would demur from agreement that that is true in the development that I have just outlined, but I would indeed like to call it a first approximation. Consciousness is not a simple matter and it should not be spoken of as if it were. Nor have I mentioned the different modes of conscious narratization such as verbal, perceptual, bodily or musical, all of which seem quite distinct with properties of their own. But it is enough, I think, to allow us to go back to the evolutionary problem as I stated it in the beginning and which has caused so much trouble in biology, psychology, and philosophy. When did all this 'inner' world begin?

We have said that consciousness is based on language. This means that generations of effort to search out the origin of consciousness in animal evolution have been incorrect and in vain. If consciousness is a group of operations learned on the basis of language, then it follows that only species with language are conscious, and that, in spite of recent questionable discussions of sign language in chimpanzees and animal communication systems, that means only in human beings, and only at some point after the evolution of language.

When did language evolve? Elsewhere I have outlined ideas of how language could have evolved from call modification, what has been called the "Wahee, Wahoo model" which is at present in competition with several others. But such theorizing points to the late Pleistocene on several grounds: (1) Such a period coincides with an evolutionary pressure for verbal communication in the hunting of large animals, (2) it coincides with the astonishing development of the particular areas of the brain involved in language, and (3), what is unique in this theory, it corresponds to the archeological record of an explosion of tool-making, for we know that language is not communication merely, but also an organ of perception, directing attention and holding attention on a particular task. This dating means that language is no older than 50,000 years, which means that consciousness developed sometime between that date and, say, the present.

It is fortunate for this problem that by 3,000 B.C., human beings have learned the remarkable ability of writing. It is therefore obvious that our first step should be to look at the early writings of mankind to see if there is evidence of an analog 'I' narratizing in a mind-space. Briefly and summarily, there is no such evidence in any writing up to approximately 1,000 B.C., and this includes the older portions of texts as well known as the Iliad.

How then did anyone decide what to do? The evidence of many sorts in all the civilizations from 9,000 B.C. up through the time of the Iliad clearly indicates that human beings heard voices, what we would regard as auditory hallucinations, voices called gods, emenating from somewhere in the brain, even as the hallucinated voices that occur in many normal people today as well as in various forms of mental illness.

This is what is called the *bicameral mind* on the metaphier of a bicameral legislature, a word incidentally which has no good translation in *German* and so may be puzzling to some of you. It simply means that human mentality at this time was in two parts, a decision making part and a follower part and neither part was conscious in a sense in which I have described consciousness. And I would like to remind you here of the rather long critique of consciousness with which I began my talk, which demonstrated that human beings could speak and understand, solve problems, and do much that we do but without being conscious. So could bicameral man. In his everyday life he was a creature of habit, but when there came some kind of problem that needed a new decision or a more complicated solution than habit could provide, that decision stress was enough to instigate an auditory hallucination, which, because such individuals had no mind-space in which to question or rebel, had to be obeyed.

I think it can easily be inferred that human beings with such a mentality had to exist in a special kind of society, one rigidly ordered in strict hierarchies with strict expectancies organized into the mind so that such hallucinations preserved the social fabric. And such was de-

finitely the case. Bicameral kingdoms were all hierarchical theocracies, with a god, often an idol, as throughout Mesopotamia, at its head from whom hallucinations seemed to come, or else, more rarely, with a human being who was divine being the head of state as in Egypt. The evidence that these early civilizations were organized by hallucinated voices called gods, I have collected elsewhere and need not go into at this time.

So far in my talk, I have discussed two of the four hypotheses of my talk. The first, which I have spent the most of my alloted time upon as being most pertinent to this symposium, is that consciousness is grounded in the ability of language to make metaphors and analogies. The second hypothesis is that in early times, there was a different non-conscious mentality called the bicameral mind. The third hypothesis is simply that one followed the other in history.

Actually, there are two possibilities here. A weak form of the theory I am presenting would state that, yes, consciousness is based on language, but instead of it being so recent, it began back at the beginning of language, perhaps even before civilization, say, about 12,000 B.C. about the time of the beginning of the bicameral mentality of hearing voices. Both systems of mind then could have gone on together, until the bicameral mind became unwieldy and was sloughed off, leaving consciousness on its own as the medium of human decisions. This is an extremely weak position because it could then explain almost anything and is almost undisprovable.

The strong form is of greater interest and is as I have stated it in introducing the concept of the bicameral mind. It sets an astonishingly recent date for the introduction into the world of this remarkable privacy of covert events which we call consciousness. The date is slightly different in different parts of the world, but in the middle East where bicameral civilization began, the date is rougly 1,000 B.C.

This dating I think can be seen in the evidence from Mesopotamia where the break-down of the bicameral mind, beginning about 1,200 B.C. is quite clear. It was due to chaotic social disorganizations, to overpopulation and probably to the success of writing in replacing the auditory mode of command. This breakdown resulted in many practices which we would now call religious which were efforts to return to the lost voices of the gods, e.g. prayer, religious worship, and particularly the many types of divination which are new ways of making decisions by supposedly returning to the directions of gods.

In Greek literature, beginning with the Linear B Tablets, going through the Iliad and the Odyssey, through the lyric and elegiac poets of the next two centuries to Solon, provides the clearest description of the breakdown of the bicameral mind and the development of the vocabulary of consciousness on the basis of metaphor. Such words as *thumos*, *phrenes*, *kardia*, *psuche* change from external objective referents to internal mental functions.

Another record of the same period is the Hebrew Testament of the Bible and it too is in agreement. The Prophets of Israel were those left-over bicameral or semi-bicameral persons who heard and could relay the voice of Yahweh with a convincing authenticity, and who were therefore prized in their societies as reaching back to the lost bicameral kingdom. The words of such prophets should be compared with the later Books of Wisdom such as Ecclesiastes, and then, of course, with the New Testament which teaches a kind of reformed Judaism for conscious people. Recently, my associate, Dr. Michael Carr, an expert in ancient Chinese texts, has demonstrated the same development of consciousness in the language of the succes-
sive sections of the *Shijing* at approximately the same time period.

But is this consciousness or the concept of consciousness? This is the well-known use-mention criticism which has been applied to Hobbes and others, as well as to the present theory. Are we not confusing here the concept of consciousness with consciousness itself? My reply is that we are fusing them, that they are the same. As Dan Dennett has pointed out in a recent discussion of this theory, there are many instances of mention and use being identical. The concept of baseball and baseball are the same thing. Or of money, or law, or good and evil. Or the concept of this Wittgenstein symposium, and this Wittgenstein symposium are the same thing.

That consciousness is based on language, that there existed a different kind of mentality called the bicameral mind, and that one followed the other are my first three hypotheses. The fourth I shall only mention briefly, as it is quite separable from the other three hypotheses and not germane to the symposium. It is a neurological model for the bicameral mind. It states that the neural substrate for the god side of the bicameral mind was what correcponds to Wernicke's Area on the right cerebral hemisphere, and that this particular location processed and learned all admonitory information in a person's life and organized it into what are heard as auditory hallucinations in order to transfer such directions to the opposite or so-called dominant hemisphere. If we use present day schizophrenics as examples of this process (schizophrenia being a partial relapse to the bicameral mind), we do find evidence that this is so. Moreover, the increasing and I must add overpopularized findings on right hemispheric function agree with this model.

In closing, I would like to point out that if we are correct that consciousness began only 3,000 years ago, this theory opens new areas for philosophical analysis as well as offering new ways to look at many ancient philosophical problems: the historical invention and rise of ethics along with its vocabulary and of law as a replacement for the voices of gods, the development of philosophy itself as a pattern of knowledge to fill the vaccuum left by the retreating gods, the historical origin of the concept of truth, and therefore of science, and of course history. I like to think that perhaps one of the reasons for the towering greatness of Greek philosophy is because the conscious world was all so new and therefore so clear, from which we could infer how our long intellectual history may work as an inhibition on our abilities and achievements.

And, of course, the greatest problem of all, the mind-body problem: the theory I have presented provides a very real solution. It restates the duality that we have been trying to understand for so long as not between mind and body, but between consciousness and everything else. This duality began only 3000 years ago. Before then, there was no consciousness-body problem whatever. And since then human nature has been ineradicably changed.

REFERENCES

Jaynes, J., *The Origin of Consciousness in the Breakdown of the Bicameral Mind*, (Boston 1976).

Jaynes, J., "The Evolution of Language in the Late Pleistocene", in: S. Harnad et al. (eds.) *Origins and Evolution of Language and Speech: Annals of the New York Academy of Sciences*, Vol. 280 (1976), p. 312.

* * *

ÜBER TAUTEGORISCHE MODELLE

Herbert Stachowiak
Universität Paderborn

1. Problemstellung

Den Begriff „tautegorisch" verwendet Friedrich Wilhelm Schelling in seinen Untersuchungen zur Mythologie als Kontrastbegriff zu „allegorisch".[1] Ist die Allegorie die versinnbildlichende Darstellung eines meist ohne explizite Kodekonvention vermittelbaren Bedeutungsinhalts, so fällt die Tautegorie mit ihrem Repräsentandum zusammen, ist mit ihm in gewisser Weise identisch. Dies, sagt Schelling, vollzöge sich in der Mythologie:

> „Die Götter sind ihr wirklich existierende Wesen, die nicht etwas anderes *sind*, etwas anderes *bedeuten*, sondern *nur* das bedeuten, was sie sind."[2]

Hiernach wäre mythologischer Sprachgebrauch durch die Extinktion einer für die menschliche Weltbegegnung fundamentalen semantischen Differenz gekennzeichnet. Er identifiziert das tautegorische „Modell" mit seinem „Original". In der Sprache der Allgemeinen Modelltheorie[3]: das Modell bildet sein Original präteritions- und abundanzfrei ab, und es ist zu ihm sowohl isomorph als auch isohyl.[4]

Was hat es mit jener tautegorischen Original-Modell-Relation des näheren auf sich? Wie läßt sie sich modelltheoretisch explizieren? Wie im Zusammenhang der Entfaltung von Philosophie verstehen? Und: Können wir heute etwas aus der Weltbegegnung des Menschen im Mythos lernen?

Dabei stellt sich natürlich unsere Analyse *über* mythologische Praxis. Sie ist *Logos* im Sinne altgriechischer Reflexion, die, wie Nestle es ausdrückt, „die Wirklichkeit in vernünftiger Rede, in begrifflichem Ausdruck, zutreffend wiederzugeben"[5] sucht. Aber diesen Logos verstehe ich hier nicht als Überwindung des Mythos. Vielmehr will ich versuchen, eine positive und für uns heute wichtige Seite des Mythologischen zurückzuholen, es womöglich neuem Verständnis aufzubereiten. Auch Schellings „Philosophie der Mythologie" verfolgte wohl dieses oder ein sehr ähnliches Ziel.[6] Mythos ist bei Schelling jener metaphysische Urgrund, aus dem sich Philosophie als „Geschichte des Selbstbewußtseins"[7] entfaltet. Schelling stand „geschichtliches Philosophieren" höher als jede geschichtslose Systematik der Vernunft. Vernunft ist je gegenwärtig und stets eingebettet in das Ganze von Vergangenheit und Zukunft. Ich stelle diese Reminiszenz ausdrücklich an den Anfang des folgenden Rekonstruktionsversuchs[8], weil ich meine, daß wir die kulturelle Dynamik, deren Resultat die heutige Welt ist, schwerlich werden verstehen können, wenn wir nicht in lernendem und vergleichend-evaluierendem Nachvollzug in den Ursprung von menschlicher Kultur überhaupt zurückgehen.

Dieses Zurückgehen bedarf der erkenntnisleitenden Zielreflexion. Meine Analyse operiert mit Modell-Original-Vergleichen auf der Grundlage von Attributensystemen und bedient sich moderner Begrifflichkeit. Mythos soll „im Lichte" solcher Begrifflichkeit in der Umkehrung seines Entfaltungsprozesses von dessen vorläufigem Endstadium her betrachtet und begriffen werden.

2. Mythologisches Original und tautegorisches Modell: Seinsweise, Struktur, Relationalität

Uns interessiert zunächst die Frage nach der Seinsweise von mythologischem Original und tautegorischem Modell. Das mythologische Original Schellings ist Teil eines kulturellen Uni-

versums, von dem wir offenlassen, aus welchen vorgängigen Ordnungsverhältnissen es entstanden ist, das jedenfalls vom Menschen erlebt wurde – vielleicht in rezenten homogenen Restkulturen partiell noch erlebt wird – als eine Gesamtheit von Ritus, Kult, Sozialgefüge und sich darin vollziehender lebensweltlicher Praxis. Das tautegorische Modell konstituiert sich demgegenüber im Individuum oder sagen wir besser: in den vielen zu der Kultur gehörenden Individuen, aus perzeptiven, emotiven und kogitativen Strukturteilen. Es gehört jedenfalls zur inneren Erlebniswelt des Menschen, der an der Kultur partizipiert. Es ist eine wesentlich psychische Entität. Das Stufenmodell der Allgemeinen Modelltheorie unterscheidet eine erste semantische Stufe von semantischen Fortsetzungsstufen. Die erste Stufe ist die der internen psychischen Gebilde. Danach kommen die Stufen der expliziten, der eigentlichen Kommunikation, der Sprache, der Schrift usw. Wir können die tautegorischen Modelle also primär der ersten semantischen Stufe zuordnen. Sie entspricht übrigens etwa der Welt 2 Poppers.

Sofern individuelle Partizipation am kulturellen Universum über Außenweltkontakte der Individuen verläuft, erfolgt sie via Zeichenträger, über das, was wir auch die Informations*substanz* nennen können. Diese entspricht Poppers Welt 1, allerdings nur in Ansehung deren Funktion als Informationsträger.[9] In der Sprache der Zeichentheorie können wir sagen, die semantische Sphäre der Kultur – Wilhelm Dupré nennt sie in seinen meisterhaften Analysen den „semantischen Raum" der Kultur[10] – vermittelt sich nur über die materiellen Medien an die Kultur-Partizipanten, durch Bedeutungsbelegung der jeweiligen materiellen Inputs gemäß einem (wir sagen heute: in der Sozialisation erlernten) Kode. – Die Seinsweise des mythologischen Originals ist schwieriger zu charakterisieren. Sicher können wir davon ausgehen, daß der „semantische Raum" der Kultur insgesamt so etwas wie eine umfassende Matrix von „Originalangeboten" an den Partizipanten bildet, und es wäre natürlich hinzuzufügen, daß dieser semantische Raum von hoher struktureller und funktioneller Komplexität ist, daß er synchronische wie diachronische Züge besitzt usw. Erinnert man dabei, daß die materiell-energetische Außenwelt für den Partizipanten dieser kulturellen Originalwelt nur Informations*träger* ist, so reduziert sich die Seinsweise jenes Universums, jener Matrix, auf etwas Konzeptionelles und Propositionales, also auf reine Bedeutungsgehalte – eben den *„semantischen* Raum" Duprés.

Wüßte man nur, was das genauer ist: Bedeutung! Es ist weder Ding noch Wort. Gleichwohl existiert es. Aber eben nicht nur mental. Und was bedeutet hier: existieren? In welchem Sinne existiert eine Norm, ein Ritus, ein Kult, ein Mythos – überhaupt ein Kulturell-Allgemeines? Von der Universalienmetaphysik über Ockhams Lösungsvorschlag[11] bis in die Gegenwart reichen die, ich darf wohl sagen: im ganzen *unbefriedigenden* Bemühungen über die Klärung des Bedeutungsbegriffs. Moore, Frege, Quine, Putnam, Davidson und viele andere haben erhebliche Anstrengungen auf diesen Problemkreis verwandt[12], dessen Aktualität auch unser Symposion zeigt. Wer die Bindestrich-Semantiken unserer Tage zusammenzählt, kommt rasch auf mehrere Dutzend. Quines semantischer Holismus verdient hier Hervorhebung, sowohl innerhalb der Semantikforschung unserer Tage als auch im rückwirkenden („reversen"[13]) Verständlichmachen der „*Kontext*semantik", die zweifellos die Symbolwelten mythologischer Kulturen beherrschte, welche Eigentümlichkeiten ihr darüber hinaus immer zuzuschreiben sind. Indes ist an Quines Holismus zu Recht dessen Naturalismus kritisiert worden.[14] Mir scheint, daß in konstruktiver Auseinandersetzung mit Quine Morton White, und zwar im Ansatz schon 1956 und dann deutlicher 1981[15], einen auch für unser Problem besonders fruchtbaren Vorschlag gemacht hat, die Schwierigkeiten nicht-holistischer semantischer Analysen holistisch – er sagt neuerdings „korporatistisch" – und pragmatisch zu überwinden (um das etwas defätistische Wort „umgehen" zu vermeiden). White gibt einem vorerst kräftigen epistemologischen Realismus Raum, beruhend auf, sagen wir, vorläufigen oder hypothetischen Evidenzen, die sich sowohl auf deskriptive wie normative Feststellungen beziehen, und legt das Gewicht auf ein je system-ganzheitliches Ausjustieren von deskriptiven und/oder normativen Widerspenstigkeiten („recalcitrant beliefs") bis hin zur pragmatisch legitimen Verwerfung fundamentaler deskriptiver Bestimmungen auf Grund der Verwerfung normativer

Konklusionen. Der Whitesche Korporatismus ist eine Lehre der *Harmonisierung* von Erfahrung und Gefühl.

Mir scheint auch für den „semantischen Raum" einer mythologischen Kultur eine solche auf das Ganze eines deskriptiv-normativen Verweisungszusamenhangs transformierte „Semantik" eine vernünftige Seinszuweisung der Originalwelt des Mythologischen. Hier allerdings ist die *dezisionäre* Form des Systemjustierens auf Grund von Widerspenstigkeiten noch eingefaltet in eine Welt selbstverständlich-traditionsbestimmter Lebenssteuerung. *Das* eben macht den Mythos – *vor* dem Logos aus. Lassen Sie mich, hoffentlich nicht *zu* spekulativ, hinzufügen, daß die mythologische Harmonie eines Seinsganzen – kritikimmun und monolitisch – vielleicht das Pendant zu jener Grunderfahrung des Menschen ist, natürlich des Menschen vor allem am Anfang seiner Kulturation, die wir die *Angst* nennen. Heidegger hat dies thematisiert[16], und ich meine, wir können seine Begründung des Seins aus der Ursprünglichkeit der Angst – die das *Nichts* offenbart[17] (als ein gleichsam Vor-Ursprüngliches, das da ist noch *vor* der Fähigkeit, *Seiendes* zu negieren) – besonders zur Bestimmung auch der Seinsweise des Mythologischen, der Welt der „Originale" tautegorischer Modellbildung heranziehen. Der Mensch hat sich *nach*-mythologisch der Angst des Nichts entledigt – kraft eben jener Negation des mehr und mehr als zu- und vorhanden Erlebten und durch ihn verfügbar Gemachten. Seine heutige „Rückkehr zur Angst" ist der Beweggrund dafür, daß ich Ihnen hier diese Gedanken über tautegorische Modelle vortrage.[18]

Im „semantischen Raum" ist angesiedelt, was weder zur physischen Welt noch zur Welt der individuellen Psyche gehört. In ihm ist angesiedelt, was *gemeinsamer* referentieller Besitz aller individueller Kulturpartizipanten ist, und dies darf eine im kulturrelativen Sinne „objektive" Welt genannt werden, die der Welt 3 Poppers gleicht. Ihre Objektivität ist die Objektivität des „kulturellen Universums", in das der junge Mensch hineinwächst, das ihn leitet, ihm Geborgenheit gibt, ihm Angst nimmt. Dabei dürfen wir offenlassen, ob – wie Popper wohl meinte[19] und wohl heute noch meint – die kultur*relativen* Drittwelten zu *einer umfassenden* Welt 3 zusammen„wachsen", die sich, vielleicht über einen kathartischen Prozeß, der nur das kulturübergreifende Bedeutsame bewahrt, als autonomes System des objektiven Geistes verewigt, so daß *diese* Welt 3 nach dem Tode auch noch des letzten Menschen in infinitum weiterexistiert.[20] Für uns hier ist, wie mir scheint, ebenso sinnvoll wie hinreichend die Unterscheidung einer je „kultur-objektiven" semantischen Sphäre, die kulturelle Kommunikation ermöglicht und in bezug auf welche der einzelne Kulturangehörende als „kulturische Monade"[21] dem lebensgemeinschaftlichen sozialkulturellen Ganzen eingefügt ist.

Ich schließe eine Bemerkung zur Struktur von mythologischem Original und tautegorischem Modell an. Die mythologischen Originale offenbaren sich in sprachlicher Gestalt, als Erzählung, eingewoben in den Alltag des Menschen jener Kultur mit seinen Riten, Geboten, Verboten, in der Vermengung von Natürlichem und Übernatürlichem. Die Entäußerungsform ist die des Zeichens, das sich der Interpretation darbietet, als techne semeiotike. Die entäußerten Inhalte sind die Wirklichkeit des Mythos, Weltanfang, Naturdeutung, Beziehung des menschlichen Seins zum göttlichen, wobei die in die Erzählung gebrachten Hauptbegriffe einerseits die allegorische Doppeldeutigkeit von Bild und Sinnbild zeigen, andererseits aber im praktischen Lebens- und Orientierungsbezug gerade aus der allegorischen Fassung heraus auf ihre *Selbst*bedeutung drängen, darauf, *das* zu meinen, was sie nominal bezeichnen: Apollon *ist* die Sonne. Phaetons Sonnenwagensturz *ist* der Sonnenuntergang im Abendrot[22]. Die Allegorie-Relate fallen schon in der mythologischen *Original*wirklichkeit semantisch zusammen, werden dem potentiellen Modellbildner als Einheit tautegorischer Identität dargeboten.

Das Zusammenfallen von Inhalt und Form, Bedeutung und Zeichen in der Wahrnehmung im Mythos betont auch Ernst Cassirer [23] unter gleichzeitiger Hervorhebung des Symbolcharakters der mythologischen Zeichen. Dieser Symbolcharakter hebe den mythischen Vorstellungsgehalt von vornherein über bloße Zeichensemantik hinaus *ins Pragmatische*.[24] Wir werden sehen, worin des näheren diese Pragmatizität besteht und wie sie den Menschen im Mythos an sich bindet, in sich einbindet. Vorerst sei hervorgehoben, was schon Vico (und in dieser Form von Cassirer merkwürdigerweise kaum aufgegriffen[25]) in seiner „Scienza Nuova"

von jener Welt mythologischen Ursprungs sagte: daß sie *archaische Kraft* ausstrahle, göttliche Ordnungsmacht, die sich als steuernde Lebensmitte der Kultur darstellt.

> Der sprachlichen Form des mythologischen Originals gemäß stelle ich mir dieses auf die Struktur eines attributiven Systems gebracht vor, lasse es also aus Individuen, Eigenschaften von Individuen und Relationen zwischen Individuen (wohl auch zwischen Eigenschaften als für sich betrachtete einstellige Prädikatoren) bestehen[26]. Diese Struktur werde auf der Metaebene unserer Betrachtung per Kodierung der formalen Elemente mit festen semantischen Referenda belegt, diese ähnlich „evidenzrealistisch" verstanden, wie ich dies für den Whiteschen Korporatismus angedeutet habe. Die mythologischen Referenda werden dem „semantischen Raum" der Kultur entnommen – genügende Separierbarkeit auf der von mir eingenommenen Ebene der Reflexion unterstellt. Kulturspezifische empirische Mythologieforschung sollte Repertoires mythologischer Originale verläßlich erarbeiten und auf explorationsfähige Systemgestalt bringen können.
>
> Entsprechend strukturiere ich die Modellseite. Auch die internen Modelle als Abbildungen mythologischer Originale betrachte ich auf der Reflexionsebene der Untersuchung als attributive Systeme. Das heißt, ich unterstelle auch dem Menschen im Mythos, und zwar auch für schriftlose Kulturen, jenes „Distinktionspotential" attributierender Modellbildung, das noch Rudiment ist, jedoch post-mythologisch zu immer vollerer Entfaltung gelangt. Dabei nehme ich für diese internen Modellbildungen diverse Struktur„liberalisierungen" an wie Fuzzy-Abgrenzungen[27], Komplexklassen- und Superzeichenbildungen[28]. Mein Attributen-„Atomismus" ist mit „sophisticated Methoden" struktuell dergestalt anzureichern, daß er Modell-Original-Vergleiche ermöglicht. Diese wiederum dürften für ein im Sinne meiner Intentionen besseres Verständnis der tautegorisch-mythologischen „Reversion" nachmythologischer, rationalisierter Modellprozesse erforderlich sein. Was ich indes in diesem kurzen Versuch hier leisten kann, geht über erste Schritte nicht hinaus.[29]

In welcher näheren Relation steht das tautegorische Modell zu seinem Original? Hierzu gehe ich von der Einbettung des Modells in pragmatische Variablen aus. Die Allgemeine Modelltheorie unterwirft jedes Modell über seinen bloßen Originalbezug hinaus dem Frageschema: Wer bildet das Modell und gegebenenfalls für wen? Für welche Zeitspanne repräsentiert das Modell sein Original? Und wozu, für welche Operationen unter welcher Zielfunktion wird das Modell gebildet? Operationalisierende Gegenstandsverfremdung jetzt auf der Metaebene unserer Analyse soll dem Gegenstand neue, bislang noch unerschlossene Information abgewinnen. Es wird sich erweisen, daß der Grenzfall der Kopierung des Originals durch das Modell nicht etwa die pragmatischen Modellvariablen funktionslos macht. Lassen Sie mich jetzt für unseren Modelltypus diese pragmatischen Variablen abchecken:

(1) Subjektivität und Benefiziarität. Eines der Basisthemen der Kulturmorphologie von Vico bis Frobenius und Toynbee ist die Sinndeutung von Geschichte und Kulturwachstum aus der Quelle des Mythos - dem Schellingschen Thema, von dem ich ausging, eng verwandt. Ursprünglich sehen wir den Menschen im Numinosen leben, wo das Göttliche in der Welt und die Welt im Göttlichen war. Diesem Monom folgte die zweite Epoche, in der Mensch und Natur, Seele und Leib auseinandertraten, dann eine dritte, in der der Mensch aufstieg zu dem abstrakten Gegensatz von Subjekt, das er selber ist, und Objekt, das die Außenwelt ist.[30] Etwa im Ausgang der ersten, auch „pneumatologisch" genannten Epoche würde ich mythologische Kultur ansiedeln. Das modellbildende Subjekt steht noch im ausgehenden Monom, hat den Weg in duale Weltverhältnisse noch kaum angetreten. Es erlebt die mythischen Zeichen des semantischen Raumes als Symptome einer Ordnung, die noch ein gesundes Ganzes ist, sich in totaler Homöostase befindet.

In einem Essay von 1983 skizziert Eugen Baer „Eine semiotische Geschichte der Symptomatologie".[31] Er charakterisiert darin „Das mythische Zeichen" als ein „totales" Zeichen, das gleichzeitig Symptom einer kultur- und lebensumfassenden Prosperität ist, aus der hinauszugeraten Krankwerden bedeutet, wobei dann individuelles Kranksein wieder aufs Ganze der Kultur wirkt. Das (modellbildende) Subjekt stand in seiner Zuständlichkeit des Gesundseins und Krankseins inmitten jener totalen Ordnung, die sich artikulierte aus Mythos – Stammeserzählungen, Helden- und Göttersagen bis schließlich hin zu den schon fast technischen Schil-

derungen Homers – und darin eingefügte rituale wie lebensweltliche Praxis. Die Symptome stimulierten unmittelbar die Therapie. Und vice versa, sagt Baer.[32] Denn das Symptom war bereits performative Aktion, auf Durchführung, Leistung, Vollzug angelegt. Es war pragmatisch im Sinne totalen Gedeihens der Kultur. Und des Vermeidens von – Angst in dieser Kultur.

So ist konkrete Subjektivität mythologisch-tautegorischer Modellbildung, wenn wir diesen Deutungen folgen, sinnvoll nur interpretierbar als Aufgehen des Subjekts in ein Anderes, als Teilnehmen, Platon spricht später in seiner artifiziellen Kulturtheorie von Teilhabe, wir sprechen heute (in einem allerdings weiterhin trivialisierenden Sinne) von Partizipation.

Die Frage nach der Benefiziarität mythologischer Modellbildung können wir schnell abhaken. Jeder in einer solchen Kultur ist bezüglich der gesamtkulturellen Wohlfahrt stellvertretend für jeden anderen da. Gerade Kultur in der Phase des Mythos ist Ausdruck des Beieinander und Miteinander.

(2) Zeitlichkeit. Modell wann? Wir heute können *bestimmen* (oder glauben, bestimmen zu können), *wann* etwas etwas anderes repräsentiert. Aber im Mythos gab es das, was ich Entscheidungsruhe nennen möchte, das Sich-Hingeben an einen als zeitlos empfundenen Geltungsanspruch[33]. Eine Zeitbestimmung der Originalrepräsentanz durch ein Modell ist unmythologisch.[34] Original und Modell sind untrennbar miteinander verbunden, das Modell taucht gleichzeitig oder mit unmerklicher Verzögerung zusammen mit seinem Original auf und ist dann von unbedingter und damit auch zeitunabhängiger Wertigkeit. Hiermit hängt zusammen, daß auch kultureller Wandel sich im Mythos bewußtlos vollzieht. Die ergodische[35] Entwicklung mythologischer Kultur ist ein „dumpfer" Prozeß. Dieser ist denkbar weit entfernt von dem, was wir (im Extrem bewußten gesamtgesellschaftlichen Tuns) „geplante Geschichte" nennen können.

(3) Intentionalität. Modell wozu? Ein operativer Modellbegriff schließt per Entscheidung gesetzte Zielfunktionen für die Modelloperationen ein. Das tautegorisch-mythologische Modell ist demgegenüber vom Ganzen der Kultur[36] her intentionalisiert, einer kollektiven Ordnung[37] eingefügt. In dieser Totalintention gehen Ansätze möglicher Partialintentionen unter. Sie werden in sie aufgehoben, wenngleich ohne derartige Partialität Kulturwandel kaum denkbar wäre, und solchen Wandel haben wir natürlich auch in den mythologischen Kulturen, lange vor dem Auftreten der Universalreligionen (dies gilt wenigstens für die heterogenen Hochkulturen, etwa Ägypten, in denen das Verhältnis von Mythos und sogenannter Realität komplexen Anpassungsprozessen unterworfen ist und Mythen Transformationen erfahren können).

Unbeschadet der *totalen* Intentionalität der modellbildenden Partizipation der Angehörigen der *homogenen* mythologischen Kulturen war diese eminent *praktisch*-zweckbestimmt. Wieder steht das Gesamtziel der Prosperität der Kultur voran, wo die Symptomatik des Mythischen als eine umfassende semiotische Ordnung für die Gesamtheit der individuellen Wohlfahrtsbedürfnisse die entscheidende Rolle spielt. Dem kulturellen Partizipanten wurde diese Ordnung in lebhaftester Weise durch die Medien ihrer so intentionalisierten Modelle fühlbar: das Hinaustreten aus der Ordnung wurde mit Krankheit und Leiden, auch mit dem Tode („Tabutod") bestraft, das ihr gemäße innere und äußere Verhalten, als Verhaltens*einheit*, mit Wohlbefinden belohnt.[38] Die von den Partizipanten der Kultur empfangenen symptomatisch-mythologischen Signalkonstellationen *taten*, *bewirkten* das, was sie bezeichneten, über die Medien der tautegorischen Modelle. Der Körper, sagt Baer, wurde so zu einer „Szene kosmischen Mysteriums".[39]

> Man ist vielleicht einen Augenblick lang geneigt, solches Ausgeliefertsein mit tierischer Instinktsteuerung über Schlüsselreiz und spannungsreduzierendem Response vergleichen zu wollen. Aber dies würde sowohl das sozial-institutionelle System menschlicher Kultur als auch deren geistig-numinoses Gefüge der inneren Daseinsermöglichung des Menschen in durchaus unzulässiger Weise ausklammern, von Sprache

als „condition humaine“ zu schweigen. Mag der *ergische* Teil menschlicher Motivdynamik den Menschen ans Tierreich knüpfen, mit seinem *engrammischen* Teil ist er dem Animalischen bereits in frühester mythologischer Verfaßtheit zweifellos in einer Weise, die solche Vergleiche nicht mehr gestattet, vollständig entwachsen. Cassirer wies übrigens auf die „Paradoxie“ hin, die er darin sah, daß sich der „Grundtrieb des Mythos“, nämlich der Trieb „zur konkret-*anschaulichen* Erfassung und Darstellung aller Daseinselemente“ gerade auf das „‚Unwirklichste‘ und Lebloseste richtet: daß das Schattenreich der Worte, der Bilder und Zeichen eine solche substantielle Gewalt über das mythische Bewußtsein gewinnt“. Den Grund hierfür sieht er darin, daß „das Dingmoment und das Bedeutungsmoment unterschiedslos ineinander aufgehen“, das heißt in der Verschmelzung von Zeichenträger und Zeichenbedeutung, und er sieht in diesem „Vereinigungszwang“ die „immanente Bedingung seiner künftigen Aufhebung“, nämlich jenes „geistigen Befreiungsprozesses, der sich im Fortschritt von der Stufe der *magisch-mythischen* Weltansicht zur eigentlich *religiösen* Weltansicht tatsächlich vollzieht“.[40]

Es wäre ein Thema für sich, und sicher ein höchst bedeutsames Thema, diese dann zunächst menschheitsentscheidende Ausdifferenzierung, die im engeren Sinne religiöse Abkopplung vom Mythischen, darzustellen und in den drei großen Verlaufsformen des Buddhismus, der jüdisch-christlich-islamischen Religionsgruppe und der innerweltlich-aufklärerischen Form der „Mythos-Überwindung“ verstehend nachzuvollziehen. Immer ist es ein Hinaustreten aus „der Einheit in das Eine oder das Andere“, wie Michael Theunissen sagt[41], und immer ist dieses Dualisieren bereits *krisis*, Krise der Prosperität der Kultureinheit, und auch wohl, wenn wir schon hier anachronistisch-vorgreifend diesen Begriff einbringen dürfen, „Krise der Macht“[42].

Die Betrachtung der Subjektivität, Zeitlichkeit und Intentionalität tautegorischer Modelle führt uns übereinstimmend immer wieder auf Vereinheitlichungen, Einheitsbildungen, und damit bestätigt sich natürlich Wohlbekanntes: die modellierenden Subjekte handeln einheitlich wie *ein* Subjekt, die Zeitspannen der Originalrepräsentation durch das Modell sind *einheitlich* zeitlicher Verfügbarkeit entzogen, die Modellintentionen, in ihrer Praxisrelevanz für jedes Mitglied der Kultur, sind zu *einer* Intention amalgamiert. Das kulturelle Universum verwirft gleichsam im Mythos die pragmatischen Differenzen, läßt individuelle Verfügbarkeit über die Form der Einbeziehung des Individuums in das verhaltensleitende Kultursystem, den Dupréschen „semantischen Raum“, nicht zu. Die tautegorische Identifikation der semantischen Originale mit ihren in der individuellen Psyche aufgebauten Modellen steht mit der Kollektivierung der Belegungen der pragmatischen Modellvariablen in einem Verhältnis wechselseitig notwendigen Bedingens. Das heißt, erst aus der Lösung von der Tautegorie, erst im Übergang zur Allegorie und damit zur Analogie, entsteht Entscheidungsfähigkeit, Wahlfreiheit des Individuums gegenüber dem Verbundsystem von Symptomen und Therapieanweisungen der Kultur, entstehen Spielräume der Symptominterpretation und damit der Therapiewahl, verlieren die Zeichen ihre Totalität und ihre Zwanghaftigkeit, wird das Ich Herr im eigenen (allerdings erst noch zu errichtenden) Haus – mit allen Nöten und Nötigungen, die solches Herrsein mit sich bringt bis zur Descartesschen Idiolatrie des „cogito ergo sum“. Solange aber Sehen und Sein *eines und dasselbe* waren, der Mensch im mythischen Zeichen Symptom *war* und es nicht etwa nur „hatte“, solange er mit dem Sehen und Sagen eines Dinges zugleich dieses Ding *war*, haben wir den Fall eines, wie es die Allgemeine Modelltheorie nennt, original-äquaten Modells oder einer Kopierung des Originals, einer strukturellen und materialen Originalidentifikation. Eugen Baer führt übrigens die geheimnisvoll-kosmische Kraft der mythologischen Symptomatologie, wenn ich ihn richtig verstehe, auf eine Art holistischen Funktionierens des „semantischen Raumes“, jetzt als Symptomraum verstanden, zurück: die Modell-Tautegorie würde paradoxerweise gerade durch die Polysemie der Symptome bewirkt, dies sei gewissermaßen der Kunstgriff der mythologischen Kultur, Weissagungen, Rituale, Trauminterpretationen usw. zu „versteinerten Konventionen“ („fossilized conventions“) und damit den einzelnen Menschen total manipulierbar zu machen – vielleicht ein wenig Orwells Big Brother vergleichbar.[43]

Soviel also hierzu. Viele Gedanken drängen sich auf. Zum Beispiel mag unser „Wahrheitssuchen“ mit der tautegorischen Modellbildung im Mythos zusammenhängen. Schelling meinte ja, Philosophie als wahrheitsfähige Wissenschaft bis in jenen noch irgendwie hinter den

monotheistischen Gottesbegriff zurückweisenden Urgrund der mythischen Einheit zurückführen zu sollen, derart, daß sich Philosophie auseinanderfalte aus dieser Einheit. Hegel sah es dann ganz anders, man könnte sagen: okzidenzialistisch, eurozentristisch, er bringt Philosophie mit Gesellschafts- und Staatsverfaßtheit in Verbindung: „In der Geschichte tritt die Philosophie da auf, wo freie Verfassungen existieren", lesen wir in „System und Geschichte der Philosophie"[44], und mit ihrem Aufstieg werde das Buch der wirklichen Welt geschrieben.[45] In Hegels rationalistischer Konstruktion, einem „Mythos aus der Retorte", manifestiert sich vielleicht, zumal wenn man die materialistische Uminterpretation dieses Modells gegenwärtigt, nach den beiden vorangegangenen Aufklärungen, der klassischen und derjenigen der sich im 17. Jahrhundert vollendenden europäischen Umschichtung, der dritte große Bruch des Menschen mit seiner mythischen Seinsverbundenheit, die für ihn *innere Daseinsermöglichung* bedeutete, mit der er hätte pfleglich umgehen sollen auch in den Stürmen der Bewußtwerdung seiner „Freiheit".

Auf Schellings Deutung der Philosophie zurückkommend, muß auch die Wissenschaftslehre des Aristoteles in neuem Licht erscheinen. Der Empirismus der Zweiten Analytiken gibt sich finitistisch, er ist in der Form logisch standardisiert, auf (apophantische) Aussagen beschränkt und enthält einige blasse methodologische Andeutungen (zum induktiven Vorgehen) – seine Nous-Verankerung jedoch reduziert ihn auf einen „formal stilisierten *Mythos*".[46] Der a-pragmatische Aspekt dieser Stilisierung zeigt sich darin, daß bei Aristoteles die mythologisch noch relevante inhaltliche Belegung der pragmatischen Modell-Variablen totaliter aufgehoben ist: das konkrete Subjekt wird zu einem nur noch formalen Subjekt *überhaupt*, es verschwindet vollständig aus dem Wissenschaftskonzept; die Geltungsdauer der Originalrepräsentation durch das Modell wird von einer konkreten Zeitspanne ins Zeitlos-Unendliche verlegt; die Zweckbindung des Modells und der Modelloperationen schließlich wird inhaltlich entleert auf absolute Zweckfreiheit. Diese formale Einebnung ist natürlich wohlzuunterscheiden von der Kollektivierung der pragmatischen Modellbestimmungen im Mythos. Was dort lebendiges, umfassendes Geschehen war, ist hier weitgehend nur lebloser Essentialismus[47], wobei sich die lebendige Prosperität der Kultur des Mythos bei Aristoteles „vereinfacht" zu einer nouetischen Intuition als erkenntnisentscheidendes „Prinzip aller Prinzipien", *dem* „Prinzip der Wissenschaft", wie uns Aristoteles belehrt, ohne diese rätselhafte Begründungsinstanz auch nur im geringsten aufhellen zu können.

Natürlich sage ich dies nicht in tadelnder Absicht. Es geht mir hier nur darum, die spezifische Transformation deutlich zu machen, die der Mythos (schon bei Homer bis ins Ironische „säkularisiert") nach der bunten Zeit der Sophisten, die ja auch einen Pragmatiker Protagoras kannte, in der attischen Spätblüte erfuhr, die Transformation in einen, sagen wir ruhig: „Wissenschaftsmythos", der erst heute, in der rezenten Epistemologie, pragmatisch reflektiert wird. Die ambivalenten Folgen dieses um die pragmatische Dimension der Erkenntnis verkürzten „Wissenschaftsmythos" sind in unserer Zeit unübersehbar geworden.

Auf die hiermit verbundene Frage, inwieweit wir bei dem Eintritt in das Reich der Freiheit das Logos-Prinzip überstrapaziert haben könnten, komme ich im dritten und letzten Teil meiner Ausführungen zurück. Hier sei vielleicht noch die Frage kurz angerührt, inwieweit wir *außerhalb* des Mythischen tautegorische Modellbildungen antreffen. Ich nehme an, in Phantasie- und Traumwelten, in denen Bedeutungsunterschiede, „ohne reaktives Verwundern" des Phantasierenden oder Träumenden, miteinander verschmelzen, sich zu Erlebniseinheiten verdichten – wahrscheinlich ohne sinnstiftende Elemente mythologischer Identifikation. Die Psychoanalyse, insbesondere psychoanalytische Traumforschung, hat z.B. das Träumen und Denken des Kindes mit dem des sogenannten Primitiven verglichen und gerade die modellierenden Identifikationsakte untersucht. Auch dichterische Phantasie mag mit Tautegorie zu tun haben. Vielleicht ist es nicht unsinnig, ein Forschungsprogramm dieses gesamten Fragenkreises zu empfehlen, in das sehr unterschiedliche Disziplinen und nicht nur die Philosophie Analysemethoden einbringen könnten. Ich für meine Person möchte hier ein starkes Informationsbedürfnis anmelden.

Entmythologisierung, später überhaupt auf Entzauberung der Welt zielend, um Max Webers Wendung wieder einmal zu strapazieren, ist Befreiung von Fesseln, die selbst zur Fessel wird. Bewußte Distinktion des Modells von seinem Original erscheint als Sündenfall. Was aus dem grenzenlosen Ganzen, dem Apeiron, hinaustritt, habe dafür Buße zu leisten, sagt Anaximander. Dem Menschen, der sich über sein Seinsganzes erhoben hat, um es zu verlassen und zu vergessen, mag sogar grausame Strafe drohen. Versuchen wir, einige der Hauptstationen des Weges jener Seinsvergessenheit in der Entfaltung von Logos und Ratio zu erinnern.

Es kann nicht verwundern, daß nach der ersten Aufklärung, der griechischen, die in der attischen Periode kulminiert, ein Totalitätsdenken, wie es im Mythos seine Wurzeln hatte, erst wieder, und zwar in enger Verbindung mit dem Christentum, in der Gnostik, im Neupythagoreismus und Neuplatonismus aufleuchtet. Das neue rationalistische Element ist unverkennbar etwa in Plotins Seinslehre mit den von einer Mitte, die Gott ist, ausgehenden Emanationsstufen.[48] Charakteristisch auch das rationalisierende Modell des Augustinus, der drei Seinsbereiche unterscheidet, das Draußen, den inneren Menschen und die Wahrheitssphäre, zu der die vernunfttätige Seele aus ihrer liebenden Beziehung zu dem personalen Gott transzendiert[49]. Einen Höhepunkt theistischer Philosophie finden wir in Avicennas Erklärungsmodell mythologisch-religiöser Gotteserkenntnis. Ausgehend von der islamischen Mystik der Seelenabspaltung entwirft er ein Modell der Dynamik von Emanation und Reversion, letztere vollzogen in asketisch-kontemplativer Lebensführung.

Mittelalterliche Mystik, im Mönchstum entstanden, häufig auch von Frauen getragen, betont mehr das Ekstatisch-Asketische, sucht Versenkung und Aufgehen in Gott – Tauler spricht direkt von der „Vergottung" im mystischen Erleben[50]. Christliche Mystik ist wohl auch als Gegenbewegung gegen intellektualisierende Theologie zu verstehen, vielleicht als ein Strebensmodus (neben dem wahrheitsphilosophischen) zurück in die Einheit des Mythischen.

Dies alles bleibt im Theologischen. Aber das Mittelalter ist ja auch die Epoche des Umbruchs, nicht nur im Sinne gesellschaftlichen Wandels, sondern auch des Wandels im philosophischen Denken. Im 13. Jahrhundert kommen empiristische Strömungen auf, der Universalienrealismus wird brüchig, alles bereitet sich auf dasjenige vor, was Panajotis Kondylis in seinem großen Aufklärungsbuch[51] dann die „Rehabilitation der Sinnlichkeit" nennt – Empirismus und Rationalismus treiben die Ich-Abspaltung vom ursprünglich mythischen Ganzen um ein jähes, großes Stück voran. Das mathematische Zeitalter kommt auf, das Zeitalter der mathematischen Modellbildungen über die paßgerecht zerschnittene Wirklichkeit, die zwar beobachtet, zuvor aber „zurechtgemacht" wird – im Hintergrund steht, zunächst noch weitgehend unbewußt, der Verfügbarkeitsgedanke. Dies schlägt auch auf Religion und Kunst über. Cusanus entwirft eine zwar auf höchste Gottesverehrung angelegte, nichtsdestoweniger menschzentrierte mathematische Theologie. Dürers Zentralperspektive ist extreme Abgrenzung des Einzel-Ichs von der Welt: auf den Fluchtpunkt ist alles hingeordnet, ihm ist alles nachgeordnet. Im Humanismus wird die Willensfreiheit nun endgültig begründet. Reformation. Neuzeitliche Naturwissenschaft, Beginn der Naturbeherrschung. Grundlegung einer neuen Mathematik, die gleichzeitig in einem Mathesis-universalis-Anspruch aufgeht. Und dann die Einmündung in Kant und die Folgen. Das sind immer neue Höhepunkte des dissezierenden, atomisierenden, das Ich, die vielen Einzel-Iche vom Ganzen abtrennenden Denkens – gegen das sich allerdings schon seit Beginn der Neuzeit zunehmend Gegenbewegungen aufrechnen lassen. Ich meine Bruno, Spinoza (die more-geometrico-Form ist seiner monistischen Philosophie im Grunde nur äußerliches Zugeständnis an eine Modeströmung seiner Zeit), den sogenannten deutschen Idealismus – auch Goethe, der das Wahre dem Göttlichen gleichsetzte, gehört zu den Antagonisten des sich etablierenden Paradigmas. Im ganzen aber geht die Separierung und Positivierung des Denkens und des diesem Denken folgenden Tuns weiter den nun einmal begonnenen Weg.

Allerdings, wer sensibel genug ist, dies zu sehen, für den beginnt sich dieser Weg *heute* in eigenartiger Vielfältigkeit aufzugabeln, und einige seiner Fortsetzungen weisen irgendwie

zurück in die Herkunftsrichtung, den Ausgangspunkt solcher Welt- und Denkentfaltung, die, wer will es leugnen, im Monströsen endete. Ja, es scheint, als bäume sich ein System, ich sage einmal in durchaus ontologisierender Absicht: ein holistisches System, *dem auch das erkennende Subjekt voll integriert ist* – als bäume sich dieses System gegen sein Zerschnittenwerden auf, gegen seine Vivisektion, als würde es nun einer tiefen gesamtsystemischen Existenznot folgend, seine Widersacher und seine Alternativen aus sich selbst gebären. Die Frühromantik kannte den Begriff des „Symphilosophierens": wirkliche Systemphilosophie mit einem, wir sagen heute, Forschungsprogramm, hier erwachsen aus dem Motiv der Menschenliebe, sollte die philosophischen Einzeldisziplinen übergreifen (ich denke hier wieder an Whites „Reunion in Philosophy"[52]) und reintegrieren. Sollte dies ein Konzept sein, das wert ist, neu durchdacht zu werden? Doch weiter in der Alternativen-Übersicht. Hier sind unter anderen sicherlich zu nennen: Schopenhauer und Nietzsche, ich meine auch: Eduard von Hartmann; in neuerer Zeit etwa Jaspers, der in seiner „Psychologie der Weltanschauungen" auch jene Denkebene, die an das Mythische und Heilige anknüpft und die er selbst ja wesentlich mitgestaltete, aus neuer Erfahrung zu reflektieren sucht. Das sind nur einige aus einer natürlich sehr viel größeren Zahl[53], und wir wissen, daß heute die Thematik „alternativen" Denkens zusammen nicht selten mit der Thematik von Mythos und Mythologie nicht nur in philosophischen Seminaren zunehmend Raum gewinnt.

> Nebenher: Ist nicht das Zustandekommen politischer Identität, *Nation* genannt, nach einer langen Phase, in der in Europa die von Papsttum und Kaisertum getragene Reichsidee herrschte, als eine wenn auch vielleicht nur „imitative" Rückkehr in jene sozialkulturelle Geborgenheit zu begreifen, die an kulturische Einheit im Mythos denken läßt? Sind nicht vielleicht die bürgerlichen Revolutionen in England und Frankreich, die den Nationalstaatsgedanken hervorriefen, zu werten als Symptome der Rückkehrtendenzen heterogenisierter Gesellschaften in „vorstaatliche" Kulturkörper mit ihren gleichsam natürlichen, werthomogenen Lebensordnungen, die auch in die Bereiche von Wirtschaft und institutioneller Daseinssicherungen ausstrahlen? Das Recht auf nationale Selbstbestimmung[54] – hat es nicht letztlich sein historisches und politisch-ethisches Fundament in einem Symbolsystem, das den Massen „das Gefühl des ‚Zu-Hause-Seins' im eigenen Land" vermitteln kann, jenseits aller Klassenschranken?[55] Aber wie könnten solche Einheiten heute noch nebeneinander Platz auf dieser Erde haben, ohne in Aporien gegenüber menschen- und völkerrechtlichen Forderungen zu geraten? Wie wären sie supranational zur „Überlebensgemeinschaft" zu integrieren auf der Grundlage von „Selbstbehauptung und Anerkennung" in einer „*globalen* Wirklichkeit der Idee der Nation", wie es Bernhard Willms fordert?[56] Ich stelle Fragen, ohne auch nur das Wagnis, sie zu beantworten, sinnvoll erwägen zu können. Vielleicht hat diese Welt die Reife einer alsbald sich vom Baum lösenden Frucht erreicht.

Einbrüche in den dominanten Rationalismus und Empirismus gibt es, wie ich hier nicht zu begründen brauche, auch in der Epistemologie. Dies beginnt schon im Konventionalismus der letzten Jahrhundertwende, bei Duhem vor allem, der zwischen zeitgemäßem Ökonomismus und unzeitgemäßem diachronisch-ganzheitlichem Denken schwankte; eine hauptsächlich über Quine führende Linie betont mehr und mehr das dynamisch-holistische Moment von Theorien- und Wissenschaftsentwicklung. Es kommt hinzu die fortschreitende rezente Historisierung von Epistemik und Epistemologie in Verbindung mit der im ganzen zweifellos fruchtbaren Kuhn-Feyerabend-Diskussion. Auch im Pragmatismus und in der Semiotik waren und sind weiterhin Züge einer Philosophie, ich möchte sagen: systemsymptomatologisch sichtbar geworden, die darauf verweisen, daß Trennungen aufgehoben, ganzheitliche Zusammenhänge zurückgeholt werden. Vielleicht sollte ich auch an die Wissenssoziologie erinnern, die in ihrer wenn auch einseitigen Art Erkenntnis umfassenderen Zusammenhängen einfügt, und erinnern auch an die Kybernetikbewegung in ihrer eigenartigen Gespaltenheit von Funktionalität und Nutzendenken einerseits, Systemdenken und einer für diese Bewegung charakteristischen Verbindung formativer, deskriptiver und normativer Elemente andererseits. Den Neopragmatismus unserer Tage, zumal in seiner systematisierten Form, thematisieren, hieße, in einen neuen Problemkreis eintreten. Dies muß ich hier zurückstellen. Rorty schließlich repräsentiert vielleicht den vorläufigen Endpunkt einer der „hermeneutisch überzogenen"

Fortsetzungslinien der gegenwärtigen pragmatischen Bewegung. Seine Gesprächs- und Bildungsphilosophie kann nur billigen, wen das Schicksal dieser Menschheitskultur[57] nichts mehr angeht.

Ist also mythisch-mythologische Welt- und Lebensorientierung gegenwärtig wieder im Kommen? Diese Frage sei zur Diskussion gestellt. Karl Acham hat im vergangenen Jahr eine bemerkenswerte Aktualität und Virulenz des Mythischen registriert, und wir alle kennen die vielfältigen obskuren Erscheinungsformen eines neuen Dranges besonders auch Jugendlicher zum Außeralltäglichen: Sektenkollektivismus, Rausch, Drogen, orgiastische Erlebnisformen unter Auslöschung der Modell-Original-Distanz, mit wortwörtlich leidenschaftlichem Bemühen, Gegenstands- und Selbstreflexion auszuschalten. Auch ziellose Apathie der Menschen unserer „Kultur" sowie die anachronistische Wiederbelebung irrationaler magisch-mythischer Rituale in entkolonialisierten Teilen der Welt gehören in diesen Formenkreis. Nun, wie dies zusammenstimmt in unserer Zeit, in der die „Aufklärung der Aufklärung" zu teils heilsamen, teils pathologischen und gefährlichen Entwicklungen geführt hat, wird wohl erst spätere kritische Kulturgeschichte überblicken können. Die Philosophie unserer Tage kann jedoch theoretisch wie hoffentlich auch praktisch dazu beitragen, den Blick für das Angemessene und Hilfreiche freizumachen, und sie sollte dies ohne Abstriche an die Forderung tun, Exaktheit der Analysen mit Ganzheitlichkeit der Gegenstandserfassung zu verbinden. An dieses, wie mir scheint, sehr zeitgemäße Anliegen mit meinem bescheidenen Beitrag zu erinnern, war der Sinn der Ausführungen, die Sie mit freundlicher Geduld für gut eine halbe Stunde über sich haben ergehen lassen.

Ich möchte diesen kleinen Versuch hier Wilhelm Dupré widmen als Dank für die von ihm erhaltenen Anregungen, ohne die ich mich, zumindest in dieser Form, kaum mit dem Problem der tautegorischen Modelle der Mythologie beschäftigt hätte.

ANMERKUNGEN

1 Schelling, F. W. J., 1973 (1856), p. 196. Vgl. daselbst Schellings Anmerkung zur Erstverwendung des Ausdrucks „tautegorisch" durch Samuel Taylor Coleridge.
2 A. a. O. (Anm. 1).
3 Stachowiak, H., 1973.
4 Zu den AMT-Termini vgl. H. Stachowiak 1973, p. 131–159, 304–333.
5 Nestle, W., 1966, p. 9.
6 Vgl. dazu K.-H. Volkmann-Schluck 1969.
7 Schelling, F. W. J., 1861, Bd. X, p. 97.
8 Rekonstruktion ist Explikation unter Einschluß von Überlegungen, die Ähnlichkeitsabbildung des Explikandums in das Explikat einer Zielfunktion zu unterwerfen.
9 Stachowiak, H., 1973, p. 199–201.
10 Dupré, W., (im Druck).
11 Vgl. E. Leinfellner und W. Leinfellner 1978, p. 64–91 (im Druck). Siehe auch den Beitrag von E. Leinfellner in diesem Band.
12 Hier der Hinweis etwa auf die von H. Delius betreute Aufsatzsammlung G. E. Moore 1969; die von G. Patzig betreute Aufsatzsammlung G. Frege 1962; W. V. Quine 1950; H. Putnam 1979; dazu die wohl „pragmatisch" zu nennende Bedeutungstheorie von D. Davidson 1980.
13 Zu diesem AMT-Terminus vgl. H. Stachowiak 1973, p. 320ff.
14 Hierzu die ausgezeichnete Analyse in H. Lauener 1982. Siehe auch den Beitrag von H. Lauener in diesem Band.
15 Vgl. M. White 1956, 1981.
16 Heidegger, M., 1943.
17 Heidegger, M., 1943, p. 13.
18 Das hier Angedeutete hat in vorzüglicher Exegese mit Bezug auf Heideggers Technikphilosophie H. Sachsse 1974 dargestellt.
19 Erste Verlautbarung auf dem Third Intern. Congr. of Logic, Methodology and Philosophy of Science, Amsterdam, August 1967.

[20] Hierzu die allerdings etwas kurzgreifende Kritik an der Dreiweltentheorie Poppers durch M. Bunge 1981, insb. p. 155ff. Popper selbst äußert sich zu diesem Punkte nicht in der wünschenswerten Klarheit. Vgl. Anm. (!) 8, p. 179, in K. R. Popper 1973.
[21] Vgl. W. Dupré 1982, p. 361f.
[22] Vgl. die in W. Burkert 1979 angeführten Beispiele.
[23] Cassirer, E., 1964 (1929), p. 72, et passim. Vgl. auch E. Cassirer 1956 (1921–1938), p. 9–25
[24] Vgl. E. Cassirer 1958 (1923), p. 8, et passim. Hierzu auch P. Verene 1982, p. 3ff.
[25] Vgl. P. Verene 1982 in Verbindung mit P. Verene 1981.
[26] Über attributive Klassen und Systeme vgl. H. Stachowiak 1973, p. 134–138, (prädikativ:) p. 305–312.
[27] Vgl. A. Kaufmann 1975.
[28] Vgl. z.B. R. Gunzenhäuser 1975, p. 124–134.
[29] In der modelltheoretischen Mythos-Analyse wären über das Sprachlich-Erzählerische hinaus gewiß auch räumlich-graphiche Äußerungen einzubeziehen, wozu bereits Ernst Cassirer in seiner „Formenlehre des Mythos" den Grund gelegt hat; vgl. E. Cassirer 1958 (1923), p. 93–128. Speziell symbolische Konfigurationen im Mythos (Figuren, „abstrakte" kultische Steingraphiken, Masken usw.) wären als mediale Modelle *zwischen* der Originalwelt, die sie repräsentieren, und den Wahrnehmungsbildern, die sie bei den Partizipanten der Kultur hervorrufen, zu untersuchen unter Herausarbeitung ihrer jeweiligen Präteritions-, Kontrastierungs- und Abundanzcharakteristik sowie der Belegungscharakteristiken für die Variablen der pragmatischen Modell-Einbettung. Zu den angedeuteten Begriffen vgl. wiederum H. Stachowiak 1973.
[30] Vgl. R. Steiner 1981 (1922), p. 36, dem ich hier nicht nur in dieser Dreiphasenunterscheidung, sondern auch in manchen seiner kraftvollen Assoziationen folgen möchte, ohne daß ich seiner mystischen Theosophie Raum gebe.
[31] Baer, E., 1983.
[32] Baer, E., 1983, p. 51.
[33] Vgl. E. Cassirer 1958 (1923), der hervorhebt, daß für die Zeitbetrachtung im Mythus „eine *absolute* Vergangenheit besteht, die als solche der weitergehenden Erklärung weder fähig noch bedürftig ist" (daselbst, p. 131).
[34] Ansätze einer Theorie des mythischen Zeitbegriffs bei E. Cassirer 1958 (1923), p. 129–169.
[35] „Ergodisch" entwickelt sich ein System, wenn es im Verlauf dieser Entwicklung von seinem Anfangszustand unabhängig wird. Zur ergodischen Entwicklung und „Selbststeuerung" von Systemen (beliebiger Art) vgl. O. Lange 1969, insbes. p. 62ff.
[36] Oder einer „mythologisch homogenen" kulturellen Einheit. In segmentären Gesellschaften z.B. können Clans je eigene Mythen oder Mythenauslegungen „haben". Das zur Modell-Intentionalität Gesagte bezieht sich dann ohne besondere Erwähnung im Text auf solche homogenen Populationen. Vgl. J. Stagl 1974.
[37] Vgl. F. Schupp 1979, p. 64, der von einer „kollektiven Axiologie" spricht.
[38] Vgl. E. Baer 1983, p. 52.
[39] Baer, E., 1983, p. 52.
[40] Cassirer, E., 1958 (1923), p. 31f.
[41] Theunissen, M., 1975.
[42] Vgl. K. Röttgers 1983, p. 52.
[43] Vgl. E. Baer 1983, p. 52f.
[44] Vgl. G. W. F. Hegel 1940, p. 227.
[45] Zum Hegel-Schelling-Verhältnis vgl. C. Cesa (im Druck).
[46] Vgl. H. Stachowiak 1971, p. 272–308. Zur Geschichte des vorsokratischen Nous-Begriffes mit besonderem Bezug auf Homer vgl. K. von Fritz 1968.
[47] Zum Aristotelischen Essentialismus vgl. E. Martens (im Druck).
[48] Zum Neuplatonismus und zur Platonrezeption durch die Väter vgl. E. von Ivánka 1964.
[49] Vgl. F. Körner (im Druck).
[50] Tauler, J., 1968, p. 162. 8ff. Ich danke diesen Hinweis U. Köpf, München.
[51] Kondylis, P., 1981.
[52] White, M., 1956.
[53] Aus der Gruppe der Gegenwartsphilosophen, die sich einen „exakten Beziehungsholismus" des Gesellschaftlichen unter den Bedingungen unserer Zeit zum Gegenstand gemacht haben, sei hier Dieter Suhr genannt, vgl. z.B. D. Suhr 1977.
[54] Hierzu zahlreiche Arbeiten Bernhard Willms', z.B. B. Willms 1983.
[55] Die zuletzt zitierte Stelle findet sich in R. Bendix 1982, p. 113. Nach Bendix hat Marx die nationalistische Idee und das Bedürfnis auch der „Unterprivilegierten" nach symbolhaft-emotionaler „Statussicherheit" innerhalb gewachsener soziokultureller Systeme fehleingeschätzt – ich würde eher sagen: ideologisch verdrängt oder bewußt heruntergespielt, um sich den Weg für die Durchsetzung *seines* Gesellschaftsmodells freizumachen.
[56] Willms, B., 1983, p. 517f. Zum Problem der Nationenbildung sowie kritisch zum Nationalstaatsgedanken K. W. Deutsch 1972.
[57] Deren Theorie – affirmativ – Arnold Gehlen geschrieben hat.

LITERATUR

Baer, E., „A Semiotic History of Symptomatology", in Eschbach, A., Trabant, J. (Ed.), *History of Semiotics* (Amsterdam – Philadelphia 1983), S. 41–66.

Bendix, R., *Freiheit und historisches Schicksal*. Heidelberger Max Weber-Vorlesungen (Frankfurt a. M. 1982).

Bunge, M., *Scientific Materialism* (Dordrecht, Holland – Boston – London, England 1981).

Burkert, W., „Mythisches Denken", in Poser, H. (Hrsg.), *Philosophie und Mythos*. Ein Kolloquium (Berlin – New York 1979), S. 16–39.

Cassirer, E., *Wesen und Wirken des Symbolbegriffs* (unveränd. Nachdr. von 4 Originalarb. 1921–1938) (Darmstadt 1956).

Cassirer, E., *Philosophie der symbolischen Formen*. Zweiter Teil: *Das mythische Denken* (unveränd. Nachdr. d. 2. Aufl. v. 1923) (Darmstadt 1958).

Cesa, C., „System und Geschichte im Spannungsfeld zwischen Schelling und Hegel", in H. Stachowiak (Hrsg.) (im Druck).

Davidson, D., „Toward a Unified Theory of Meaning and Action", *Grazer Philosophische Studien* 11 (1980), S. 1–12.

Deutsch, K. W., *Nationenbildung – Nationalstaat – Integration* (Düsseldorf 1972).

Dupré, W., „Tradition und Aneignung", in Nagl-Docekal, H. (Hrsg.), *Überlieferung und Aufgabe*. Festschrift für Erich Heintel zum 70. Geburtstag, 2. Teilband (Wien 1982), S. 355–372.

Dupré, W., „Das Pragma in schriftlosen Kulturen", in Stachowiak, H. (Hrsg.) (im Druck).

Frege, G., *Funktion, Begriff, Bedeutung*. Fünf logische Studien, Hrsg. u. eingel. v. G. Patzig (Göttingen 1962).

Fritz, K. von, „Die Rolle des NOYΣ", in Gadamer, H. (Hrsg.), *Um die Begriffswelt der Vorsokratiker* (Darmstadt 1968), S. 246–363.

Gunzenhäuser, R., *Maß und Information als ästhetische Kategorien*. Einführung in die ästhetische Theorie G. D. Birkhoffs und die Informationsästhetik (Baden-Baden ^{2}1975).

Hegel, G. W. F.,*System und Geschichte der Philosophie*. Hrsg. v. J. Hoffmeister (Hamburg 1940).

Heidegger, M., *Was ist Metaphysik*? (Frankfurt a. M. 1943).

Ivánka, E. von, *Plato Christianus*. Übernahme und Umgestaltung des Platonismus durch die Väter (Einsiedeln 1964).

Kaufmann, A., *Introduction to the Theory of Fuzzy Subsets* (New York, N.Y. – San Francisco, Cal. – London 1975).

Kondylis, P., *Die Aufklärung im Rahmen des neuzeitlichen Rationalismus* (Stuttgart 1981).

Körner, F., „Existentielle Pragmatik aus transzendentaler Immanenz als innere Tragkraft letzter Wahrheitssuche am geistesgeschichtlichen Beginn abendländischen Rationalismus", in Stachowiak, H. (Hrsg.) (im Druck).

Lange, O., *Ganzheit und Entwicklung in kybernetischer Sicht* (aus d. Poln.) (Berlin 1969).

Lauener, H., *Willard V. Quine* (München 1982).

Leinfellner, E., Leinfellner, W., *Ontologie, Systemtheorie und Semantik* (Berlin 1978).

Leinfellner, E., Leinfellner, W., „Ockhams Semantik und Pragmatik", in Stachowiak, H. (Hrsg.) (im Druck).

Martens, E., „Platonischer Pragmatismus und Aristotelischer Essentialismus", in Stachowiak, H. (Hrsg.) (im Druck).

Moore, G. E., *Eine Verteidigung des Common Sense*. Fünf Aufsätze (aus d. Engl.). Eingel. v. H. Delius (Frankfurt a. M. 1969).

Nestle, W., *Vom Mythos zum Logos*. Die Selbstentfaltung des griechischen Denkens von Homer bis auf die Sophistik und Sokrates (Aalen 1966).

Popper, K. R., *Objektive Erkenntnis*. Ein evolutionärer Entwurf (aus d. Engl.) (Hamburg 1973).

Putnam, H., *Die Bedeutung von „Bedeutung"* (aus d. Amer.). Hrsg. u. Übers. v. W. Spohn (Frankfurt a. M. 1979).

Quine, W. V., *Methods of Logic* (New York, N. Y. 1950).

Röttgers, K., *Texte und Menschen* (Würzburg 1983).

Sachsse, H., „Was ist Metaphysik? Überlegungen zur Freiburger Antrittsvorlesung von Martin Heidegger und ein Exkurs über seine Frage nach der Technik", *Zs. f. Philos. Forschung* 28.1 (1974), S. 67–93.

Schelling, F. W. J., *Sämtliche Werke* (14 Bde.) (Stuttgart – Augsburg 1856–1861).

Schelling, F. W. J., *Philosophie der Mythologie* Bd. I: *Einleitung in die Philosophie der Mythologie* (unveränd. Nachdr. d. v. K. F. A. Schelling aus d. Nachlaß hrsg. Ausg. v. 1856) (Darmstadt 1973).

Schupp F., „Mythos und Religion: der Spielraum der Ordnung", in Poser, H. (Hrsg.), *Philosophie im Mythos*. Ein Kolloquium (Berlin – New York 1979), S. 59–74.

Stachowiak, H., *Allgemeine Modelltheorie* (Wien – New York 1973).

Stachowiak, H., (Hrsg.), *Pragmatik*, Bd. I: *Pragmatisches Denken von den Ursprüngen bis zum 18. Jahrhundert* (Hamburg (im Druck)).

Stagl, J., *Die Morphologie segmentärer Gesellschaften*, Dargestellt am Beispiel des Hochlandes von Neuguinea (Meisenheim am Glan 1974).

Steiner, R., *Der Entstehungsmoment der Naturwissenschaft in der Weltgeschichte* (Vortragskurs, geh. in Dornach 1922/23) (Dornach 1981).
Suhr, D., *Die kognitiv-praktische Situation*. Fundamentierungsprobleme in praktischer Philosophie, Sozialtechnik und Jurisprudenz (Berlin 1977).
Tauler, J., *Die Predigten Taulers*. Hrsg. v. F. Vetter (= Deutsche Texte des Mittelalters 11). Nachdr. (Dublin – Zürich 1968).
Theunissen, M., „Krise der Macht", in: *Referate des X. Internat. Hegel-Kongresses in Moskau* (Köln 1975 (Hegel-Jahrb. 1974)), S. 318–329.
Verene, P., *Vicos Science of Imagination* (Ithaca, N. Y. 1981).
Verene, P., „Cassirer's Philosophy of Culture", *Intern. Philos. Qu.* 22 (1982), S. 133–144.
Volkmann-Schluck, K.-H., *Mythos und Logos*. Interpretationen zu Schellings Philosophie der Mythologie (Berlin 1969).
White, M., *Toward Reunion in Philosophy* (Cambridge, Mass. 1956).
White, M., *What Is and What Ought To Be Done* (New York, N. Y. – Oxford 1981).
Willms, B., „Weltbürgerkrieg und Nationalstaat. Thomas Hobbes, Friedrich Meinecke und die Möglichkeit der Geschichtsphilosophie im 20. Jahrhundert", *Der Staat* 4 (1983), S. 499–519.

* * *

A LOGICAL CHARACTERIZATION OF THE PSYCHOLOGICAL

Roderick M. Chisholm
Brown University, Providence, R. I.

Introduction

If we are going to talk about the philosophy of mind and the "mind-body problem," we should have a way of characterizing the psychological *logically* and without making use of any psychological concepts.[1]

I will assume that, for each of us, there is a set of psychological properties which are such that we can know directly and immediately that we have those properties. And I will also assume that we can reflect upon these properties and in so doing come to know something about their general structure. Thus I now know directly and immediately such things as these: I hope we will have a useful discussion; I look forward to walking in the hills; I believe I have been here before; I seem to see something red.

If I may speak in the manner of G. E. Moore, I would also say, not only that it is evident to me that I have such properties, but also that it is evident to you that you have such properties. I know, of course, that today there are many philosophers who deny these things. But their denial, in every case, is based upon data far less obvious than the facts they profess to deny.

The properties I am concerned with, then, are properties that we might cail "Cartesian" —such properties as: thinking; wishing; desiring; believing; sensing. The properties are *psychological* in that, if you know of a person that he has such a property, then you know something that relates him to the property of thinking.

All the properties I have referred to are *purely* psychological—they tell us *only* about thinking. They tell us nothing about the person's bodily state, nothing about his past, nothing about his future, and nothing about any other individual.

Some Methodological Comments

I have said that psychological properties are properties to which we have a certain privileged access. But I do not intend to *define* the psychological this way, since privileged access is itself a psychological concept.

In proposing a *logical* characterization of such properties, I will make use of the following undefined concepts: that of *exemplification*, as in "x exemplifies y"; that of *de re necessity*, as in "x is necessarily such that it exemplifies y"; that of an *individual thing*, and those of possibility, potentiality and disposition. I will sometimes use "has" as short for "exemplifies"; and I will use "x is possibly such that it is so-and-so" and "x can be so-and-so" as short for "x is not necessarily such that it is not so-and-so." And I will rely upon distinctions of *tense*.

I will now make explicit certain metaphysical assumptions. (1) One is that everything is necessarily such that it undergoes change. This is the assumption that is sometimes put by saying "everything is necessarily in time." (2) Another assumption is that it is logically possible that there is only one individual thing. One might put this by saying "there is a logically possible world in which there is only one individual thing." There may well be reasons for supposing that, if there are individual things, then there are many of them. But it is at least *logically* possible that there is only one. (3) I will also make a mereological assumption—an assumption about wholes and parts. It is this: for any two individuals that have no parts in common, there is another thing that has those two individuals as proper parts. Making this assumption is tan-

tamount to countenancing the being of scattered objects and other "mere aggregates." (4) The final metaphysical assumption is that of extreme realism—Platonism, in the contemporary sense of that word.

My general strategy is as follows. I will begin with some examples of purely psychological properties. I will single out certain logical features that are common and peculiar to our examples. These features will enable us to define the concept of *that which is purely psychological*. We may then define *the psychological* as that which includes the purely psychological.

When we do philosophy, we have to begin somewhere. Perhaps you don't begin at the same place that I do. But this does not mean that one of us is wrong. If we are philosophically sophisticated, we will know that the important thing is to be clear about what it is that one assumes at the outset and about what follows from these assumptions.

Changeable Properties of Individuals

It will be useful to take the term "property" in a somewhat restricted sense.

D1 P is a property =Df P is possibly such that there is something that exemplifies it

This definition, like those that follow, is in the present tense. I will assume that the past and future versions of these definitions are obvious.

Among properties, in the present sense of the word, are: being green; being heavy; being a table; being a horse; being a unicorn; being round; being square. But, although being round and being square are properties, in our present sense, being round *and* square is not a property. (If you wish to call *being a round square* a property, then you should replace "property" in my definition by some other word—say "attribute" or "determination.")

Let us now introduce the concept, *being a changeable property of an individual*. This concept will give us our first mark of the purely psychological.

Perhaps it is uncontroversial that purely psychological properties (feeling sad, hoping for rain, believing that all men are mortal . . .) are "changeable properties of individuals" in the following sense:

D2 P is a changeable property of individuals =Df P is necessarily such that only individuals have it; P is possibly such that many individuals have it; and whatever can have P can also cease to have it and then have it again, any number of times

(I would remind you that the possibility and necessity of which we here speak are *logical possibility* and *logical necessity*.)

Thinking is a changeable property of individuals, but *being extended* is not. And, although thinking is a changeable property of individuals, the property of *having thought* is not; once you take it on, you will have it as long as you exist. And the property of *being such that one is going to think* is not a changeable property; once you lose it, you can never take it on again.

Here, then, is the first of our three marks of the purely psychological. A purely psychological property is a changeable property of individuals.

Internal Properties

The second mark of the purely psychological is this: Every purely psychological property is *internal*. What may be new, I think, is that this is a *logical* characterization of internality—a characterization which does not make use of metaphors and one which does not appeal to "privileged access" or to "inner perception."

Let us begin by singling out a certain logical relation that may hold between properties. I will call this relation—somewhat arbitrarily—the relation of property *inclusion* and define it as follows:

D3 P includes Q =Df P is necessarily such that, for every x, if x exemplifies P, then x exemplifies Q

The property *walking* includes the property *having limbs*. And the property *walking swiftly* includes the property *walking*. But since our definitions are in the present tense, we cannot say that the properties of *having walked* and of *being such as to be going to walk* include the property *walking*. For a thing that is not walking may yet be such that it did walk or it is going to walk.

The properties of *having walked* and of *being such as to be going to walk* may be said to *project* the property of *walking*. The relevant sense of "project" is this:

D4 P projects Q =Df P is necessarily such that, if it is exemplified, then either (a) Q is exemplified or (b) Q was exemplified or (c) Q will be exemplified

The property of being a bride-to-be projects the property of being a husband; so, too, does the property of being a widow. The property of having been thought to be a thief projects the property of being thought to be a thief as well as that of thinking, but it does not project the property of being a thief.

It is obvious that, if a property P includes a property Q, then P projects Q—and not conversely.

We may now characterize internal properties in terms of the kinds of properties they project.

Let us first note that there are properties that are projected by *every* property. These are the properties that certain abstract objects have necessarily. Examples are: being the successor of 7; being the property blue; being self-identical; and being such that 7 and 5 are 12. Every property is necessarily such that, if it is exemplified, then these necessary properties are exemplified.

Now we may note a peculiarity of properties that are internal. Except for the necessary properties just referred to, the only properties that an internal property projects are those properties that it includes. Hence we may define *internal property* this way:

D5 P is an internal property =Df For every Q, if P projects Q and if only individual things can have Q, then P includes Q

Feeling sad, hoping for rain, believing that all men are mortal, along with the other "purely psychological properties" with which we began, satisfy the terms of this definition of internal property.

If a property is internal, then, it is possibly such that only one individual has it; indeed, it is possibly such that there will have been only one individual and that that individual has it. Thus the property *being sad* is possibly such that there is an individual that has it and there are no other individuals. It does not require the existence of any individual other than its own bearer. We may express this point by saying that internal properties are not "other implying".

That internal properties are not other-implying follows from the definition and our assumption that it is logically possible that there is only one individual thing. Consider a property that *is* other-implying—say, the property of being one of two individuals. This property projects a property it does not include (that of being an aggregate), and therefore it is not internal.

If we take "red" in its physical sense, we may say that *being red* is other-implying. For it projects but does not include the property of *being a proper part*. A thing cannot be red unless there are things that are proper parts of other things. But a thing may be red without itself being a proper part of any other thing. *Locomotion* is also other-implying; for it implies that there are two individual things changing their spatial relations to each other. And so, too, for the other properties that Locke had called "primary qualities"—"solidity, extension, figure, motion or rest and number."[2] And *weight* is also other-implying (for to say that a body has weight is to say that it stands in certain relations to some other body or bodies).

So far, then, we have two marks of the psychological. A psychological property is a changeable property of individuals and it is internal. If we turn now to nonpsychological properties,

we are likely to find that the first ones to come to mind—say, *being 6 feet long* and *being such as to have 3 dimensions*—are such that either they are not internal or they are not changeable properties of individuals. We may be tempted to say, therefore, that *only* psychological properties have these features. But this would be a mistake. For there are nonpsychological properties that are changeable properties of individuals and that satisfy our definition of internality.

Consider that disjunctive property that all sad people and islands have in common—the property of being either sad or entirely surrounded by water. This both is a changeable property and is internal by our definition.

Hence we must look further if we are to find a logical *definition* of the psychological.

In the early part of the present century many philosophers spoke of properties that are "qualitative." This concept gives us a third mark of the psychological.

Purely Qualitative Properties

If a property is is purely qualitative, then it is not equivalent to a mere potency, disposition or possibility, But this is only a negative characterization. To characterize the concept more positively, let us introduce the following useful abbreviations:

D6 P falls under Q =Df P includes Q; and Q does not include P

Being a dog falls under being an animal, and believing that there are dogs falls under believing. (Please note that I am now using the expression "falls under" to refer to a relation that holds only between properties or relations. Using the expression in this sense, we cannot say that a thing "falls under" the properties that it exemplifies.)

D7 P falls between Q and R =Df Q and R are such that (a) P falls under one and (b) the other falls under P

Being a dog falls between being a wolfhound and being an animal, and believing that there are dogs falls between believing and believing that there are dogs and cats.

Some properties are such that they may be said to "*fall densely* under other properties. The concept may be illustrated by reference to the properties:

(P) being between 5 and 7 feet tall
(Q) being 6 feet tall

Consider any property R that falls between these properties P and Q. We might have, for example:

(P) being between 5 and 7 feet tall
(R) being between 5.1 and 6.9 feet tall
(Q) being 6 feet tall

For any property R, which thus falls between P and Q, we may find a property that falls between Q and R, and we may find a property that falls between P and R. The property P, then, "falls densely" under Q. The concept may be defined this way:

D8 P falls densely under Q =Df P falls under Q; and for every R, if R falls between P and Q, then (a) something falls between P and R and (b) something falls between Q and R

The word "dense" is appropriate, for such properties are related to each other in a way in which fractions are related to each other: I am thinking of the fact that between any two fractions there are indefinitely many other fractions.

We may note that, although *being 6 feet tall* falls densely under *some* properties, it falls under *other* properties without falling densely under those other properties. Thus it falls, but not densely, under *being possibly 6 feet tall*; there are no properties that fall between these properties.

These considerations lead us to the concept of the purely qualitative. Consider the property of *thinking*—the property of *being conscious*. Shall we say that the property of thinking is a property that does not fall densely under any property? This would not be accurate. For thinking falls densely under this property: *being a member of a class which is such that more than half of its members are thinking*. Between this latter property and thinking, we find, for example, *being a member of a class which is such that more than three quarters of its members are thinking*. It is clear that *thinking* falls densely under such properties. But these properties, unlike thinking, are not internal. For they project a property they do not include—the property of thinking.

Let us, then, define "purely qualitative property" this way:

D9 P is purely qualitative =Df P is a property which is such that (a) nothing can have P necessarily, (b) P does not fall densely under any property that is internal, and (c) P is not a mere potency, disposition or possibility

Thinking is qualitative by this definition; so, too, for *being able to think*. (We should note, in passing, that there is a significant difference between the two purely qualitative properties—thinking and being able to think. Unlike thinking, *being able to think* is not one of those properties which are such that, if you have it, then you can know directly and immediately that you have it. But it is such that, if you have it and if you *raise the question* whether you have it, then you *can* know directly and immediately that you have it.)

The following subspecies of thinking are also purely qualitative: believing, liking, disliking, sensing visually, sensing auditorily, sensing tactually. But it is problematic whether *all* subspecies of thinking are purely qualitative.

Does *sensing redly*, for example, fall densely under *sensing visually*? (The question may be put in a different terminology: Does *having a red sensation* falls densely under *having a visual sensation*? We may note that *having a visual sensation* does not fall densely under anything.)

Again, does *believing that there are dogs and cats* fall densely under *believing that there are dogs*? One might be tempted to say that it does—on the ground that *believing that there are dogs and other animals* falls between these things. But does *believing that there are dogs and cats* include *believing that there are dogs and other animals*?

Whether or not such properties are themselves purely qualitative, each of them *includes* a property that is purely qualitative.

A Definition of the Psychological

Shall we now say that a psychological property is a purely qualitative property? This would be too narrow. Consider the property of judging truly that there is someone who is ill. Surely this is psychological. Yet it projects the property of being ill without including that property. Hence it is not internal, and therefore it is not purely qualitative.

Let us say rather this:

D10 P is psychological =Df P is a property; and P includes a property that is (i) a changeable property of individuals; (ii) internal, and (iii) purely qualitative.

Here, then, we have our logical characterization of the psychological.

Some Problematic Cases

I will speak here only of two types of problematic case—*obligations* and *semantic* properties. In the past, these two types of property have served as obstacles in the way of finding a logical criterion of intentionality.[3]

Does *being obligated* project *thinking* and thus project a property it does not include? One might argue: "Surely no one can be under an obligation without having been conscious at some time or other." If being obligated does thus project thinking, it projects a property it does not include, and therefore it is not an internal property. But one may also argue that being obligated does *not* thus project thinking. "Consider a person who has never been conscious. Isn't such a person also under obligation? After all, he, too, has the obligation—the duty—not to commit murder." But what is it thus to have the duty not to commit murder? It is to have the property of being such as *not* to have the *right* to commit murder. But this property is not restricted to individuals. And therefore it is not a "changeable property of individuals."

Consider now the *semantic* property: having dogness (the property of being a dog) as part of its meaning. This is a property of the English words "dog" and "bulldog." In what respects does this property differ from psychological properties?

There are two ways of conceiving the property in question: (a) one may conceive it as a property of certain abstract objects; or (b) one may conceive it as a changeable property of individuals. In the former case, it will not satisfy the conditions of the definition of psychological property. In the latter case, it may be thought of as a property of such things as the written words, "dog", "chien", "Hund"; hence it would not be an internal property—for its exemplification would require that there be more than one individual thing.

ENDNOTES

1 In "On the Nature of the Psychological," *Philosophical Studies*, Vol. 43 (1983), pp. 155–164, I presented a definition of psychological property; but the definition was not purely logical since it made use of the psychological expression "to conceive."

2 See Locke's *Essay Concerning Human Understanding*, Book II, Ch. 8, Section 10.

3 See the article "Intentionality" in Paul Edwards (ed.), *The Encyclopedia of Philosophy* (New York 1967), Vol. IV, pp. 201–204.

* * *

ZEITLICHKEIT DES BEWUSSTSEINS UND SYNTHETISCHES URTEIL A PRIORI

Konrad Cramer
Universität Göttingen

Der Grundgedanke von Kants nicht-empirischer Theorie der empirischen Erkenntnis läßt sich doxographisch in wenigen Sätzen zusammenfassen. (1) Wenn ein Subjekt, dem Sinnesdaten auf Grund seines Vermögens, Vorstellungen zu ‚empfangen', gegeben werden, sich diese als seine eigenen soll zuschreiben können, dann muß es sich der Identität seiner selbst mit Bezug auf solche Daten bewußt sein können. (2) Seiner Identität mit Bezug auf gegebene Vorstellungen kann sich ein solches Subjekt nur dann bewußt sein, wenn es diese Vorstellungen zur Einheit eines Bewußtseins verbindet. (3) Das auf einer Synthesis gegebener Vorstellungen beruhende Bewußtsein von der Einheit des Bewußtseins kann selber keine gegebene Vorstellung sein. Es beruht vielmehr auf einem ‚Aktus der Spontaneität der Vorstellungskraft', der sich in der Funktion des Urteils bekundet, durch welches ein Mannigfaltiges gegebener Vorstellungen zur Einheit eines Bewußtseins gebracht wird. (4) Die Analyse des durch die Funktion des Urteils erzeugten Einheitssinnes in der Verbindung gegebener Vorstellungen zeigt, daß das Subjekt über gewisse Grundbegriffe, die Kategorien, verfügen muß, in denen eine solche Einheit gedacht wird. (5) Sie zeigt weiter, daß sich ein Subjekt der Identität seiner selbst mit Bezug auf gegebene Vorstellungen dann und nur dann bewußt sein kann, wenn es durch den Gebrauch von Kategorien die Inhalte dieser Vorstellungen als Bestimmungen von etwas auffaßt, das von dem bloßen Vorgestelltsein dieser Vorstellungen unterschieden ist. (6) Dieses von den jeweiligen Vorstellungszuständen des Subjekts Unterschiedene ist der Gegenstand oder das Objekt seiner Vorstellungen.

Es ist insbesondere die Version, die Kant der ‚Transzendentalen Deduktion der Kategorien' in der 2. Auflage der ‚Kritik der reinen Vernunft' (B) gegeben hat, welche den in (1)–(6) skizzierten Zusammenhang zwischen dem Bewußtsein der Identität eines Subjekts mit Bezug auf eine Mannigfaltigkeit von Vorstellungen, die ihm als sinnliche gegeben sind, und der Beziehung dieser Vorstellungen auf einen Gegenstand der Erkenntnis herzustellen unternimmt. Zwar ist die Leistungskraft der Argumentationsschritte, in denen dies geschieht, bis heute noch nicht vollständig aufgeschlossen; und diese Schritte selber sind genau genommen noch nicht einmal mit Eindeutigkeit identifiziert.[1] Sicher ist jedoch, daß diejenige Überlegung Kants, welche den Kategorien die Funktion zuschreibt, das Mannigfaltige einer Anschauung, die ein auf gegebene Vorstellungen angewiesenes Subjekt die seinige nennt, „durch die Synthesis des Verstandes als zur notwendigen Einheit des Selbstbewußtsein gehörig" (B 144) vorzustellen, noch ganz davon absieht, *auf welche Weise*, d.h. unter welchen Formen der Sinnlichkeit einem Subjekt ein Mannigfaltiges sinnlicher Vorstellungen gegeben wird. Gefordert ist hier nur, daß ein Subjekt das Vermögen besitzt, *überhaupt* sinnlich anzuschauen, d.h. mit Bezug auf gewisse Vorstellungen, die es sich als die seinen zuschreibt, das Bewußtsein davon zu haben, daß sie nicht von ihm selber erzeugt, sondern ihm „noch vor der Synthesis des Verstandes und unabhängig von ihr" gegeben sind – „wie aber, bleibt hier unbestimmt." (B 145)

Wenn demnach in strenger Allgemeinheit gilt, daß ein Subjekt sich gegebene Vorstellungen der Sinnlichkeit nur dann als seine Vorstellungen zuschreiben kann, wenn es die in ihnen präsentierten Inhalte auf ein von ihnen unterschiedenes Objekt zu beziehen vermag, dann muß dies *eo ipso* auch für ein solches Subjekt gelten, dessen für seine Sinnlichkeit spezifische Anschauungsform die *Zeit* ist. Mit Rücksicht auf die Zeitlichkeit des Gegebenseins von Vorstellungen ergibt sich daher für Kant weder ein neues noch ein zusätzliches Deduktionsproblem, sondern nur die Frage, welche *spezifische* Bedeutung den Kategorien als denjenigen

Begriffen a priori, in denen eine synthetische und objektive Einheit gegebener Vorstellungen der Sinnlichkeit im Modus des ‚überhaupt' gedacht wird, in ihrer Anwendung auf die unsere Sinnlichkeit kennzeichnende Auschauungsform – die Zeit – zuwächst. Diese Frage versucht Kant in den Kapiteln über den ‚Schematismus der reinen Verstandesbegriffe' und über das ‚System der Grundsätze des reinen Verstandes' zu beantworten. In der Auszeichnung und Begründung eines solchen Systems will Kant den Nachweis führen, daß für ein Subjekt, dem Vorstellungen im *Wechsel* gegeben sind, ein genau bestimmter Inbegriff von synthetischen Urteilen a priori in Geltung stehen muß, deren Aussagegehalt die Beziehung zwischen der Selbstzuschreibung solcher Vorstellungen von seiten dieses Subjekts und der Beziehung der in ihnen präsentierten Inhalte auf einen Gegenstand der Erkenntnis in der Zeit (und, wie zu ergänzen ist, im Raum) näher, d.h. mit Rücksicht auf die Anschauungsform der Zeit (und des Raums) bestimmt. Solche Urteile sind für ein zeitlich anschauendes Subjekt insofern ‚objektkonstituierende' Urteile, als es über sie verfügen können muß, um die ihm auf spezifische Weise gegebenen Vorstellungen seiner Sinnlichkeit in entsprechend spezifisch bestimmter Weise auf einen Gegenstand einer *ihm* möglichen Erfahrung beziehen zu können. Diejenigen synthetischen Urteile a priori dieses Systems, die auf das ‚Dasein' der Gegenstände einer einem solchen Subjekt möglichen Erfahrung gehen, faßt Kant unter dem Titel der ‚Analogien der Erfahrung' zusammen.

Der erste Halbsatz der 1. Analogie der Erfahrung lautet: „Bei allem Wechsel der Erscheinungen beharrt die Substanz." (B 224) Der Beweis dieses synthetischen Urteils a priori hat nach Kants eigenen methodologischen Bemerkungen über die Beweisaufgabe ‚transzendentaler Sätze' darzutun, daß seinem Subjektbegriff (‚Wechsel der Erscheinungen') durch die synthetische Beziehung des Prädikatbegriffs (‚Substanz') in der temporalen Bestimmtheit von ‚Beharrlichkeit' auf ihn die Bedeutung des Begriffs von einem *Objekt* einer einem zeitlich anschauenden Subjekt möglichen Erfahrung allererst zuwächst.

Inwiefern, so ist zunächst zu fragen, ist der erste Halbsatz der 1. Analogie ein *synthetisches* Urteil? Offensichtlich ist er dies nur dann, wenn der Begriff ‚Wechsel der Erscheinungen', nach seinem logischen Inhalt betrachtet, noch nicht diejenigen Konnotationen des Begriffs ‚Erscheinung' aufweist, denen gemäß nach Kants Lehre vom transzendentalen Idealismus von Zeit (und Raum) gerade die *Objekte* einer uns möglichen Erfahrung ‚bloße Erscheinungen' und nicht ‚Dinge an sich' sind. Denn der Begriff der Erscheinung als Begriff von einem Objekt einer uns möglichen Erfahrung enthält den Begriff der Substanz in seiner näheren Bestimmung durch den Zeitmodus der Beharrlichkeit gerade *analytisch* in sich. Erscheinungen als Objekte einer uns möglichen Erfahrung sind Entitäten, die nach dem kategorialen Schema der Beharrlichkeit von etwas im Wechsel oder im Zugleichsein seiner Zustände gedacht werden müssen, weil eben dieser Gedanke ihren Objektcharakter definiert. *Daß* dies der Fall ist, sagt der erste Halbsatz der 1. Analogie jedoch gerade aus, und zwar – der Behauptung nach – nicht auf analytische, sondern auf synthetische Weise. Soll diese Behauptung zutreffen, kann der Subjektbegriff der Aussage nur der Begriff vom Wechsel der Erscheinungen als *bloßer Vorstellungen* sein, die einem Subjekt in der Form der Zeit gegeben werden, also nicht selber schon der Begriff von der gerechtfertigten Beziehung solcher Vorstellungen auf ein von ihnen unterschiedenes Objekt derselben.

Kants Behauptung ist ferner, daß der erste Halbsatz der 1. Analogie ein Urteil *a priori* ist. Die in ihm zu denkende Synthesis von Subjekt- und Prädikatbegriff erfolgt nicht auf Grund der Erfahrung von Objekten in der Zeit und im Raum, sondern benennt eine ‚Bedingung der Möglichkeit' eben dieser Erfahrung. Genau dann, wenn dies der Fall ist, ist die Aussage kein empirisches Urteil, sondern ein Urteil, das unabhängig von aller wirklich gemachten besonderen Erfahrung von Objekten in Zeit und Raum, mithin a priori, gilt.

Diese Kennzeichnung wirft jedoch ein eigenes Problem auf. „Empirische Begriffe, imgleichen das, worauf sie sich gründen, die empirische Anschauung, können keinen synthetischen Satz geben, als nur einen solchen, der auch bloß empirisch, d.i. ein Erfahrungssatz ist, mithin niemals Notwendigkeit und absolute Allgemeinheit enthalten kann." (B 64) Diese Feststellung Kants leuchtet in der Tat ein. *Empirische* Begriffe können zwar in synthetischen Urtei-

len, aber nicht in synthetischen Urteilen a priori Verwendung finden, weil Urteile, in denen sie als Termini auftreten, nicht unabhängig von derjenigen empirischen Anschauung sind, deren allgemeine Vorstellung solche Begriffe sind. Begriffe, die in synthetischen Urteilen a priori Verwendung finden können sollen, müssen ihrerseits von aller Erfahrung, mithin auch von dem, auf was sich eine bestimmte Erfahrung gründet, der empirischen Anschauung, unabhängig sein. Unter dieser Kautele ergibt sich jedoch für den Subjektbegriff der 1. Analogie der Erfahrung, den Begriff ‚Wechsel der Erscheinungen', daß er der von Kant selber formulierten Bedingung für die Apriorität von Begriffen nicht zu genügen scheint. Denn nur *reine* Anschauungen oder *reine* Begriffe sollen a priori möglich sein. (Vgl. B 75) ‚Rein' aber sind Anschauungen oder Begriffe, wenn in ihren sinnlichen oder logischen Inhalt keinerlei Bezugnahme auf empirisch gegebene Daten unserer Sinnlichkeit eingeht. Dieser Forderung genügt der Begriff des ‚Wechsels der Erscheinungen' nicht.

Der Nachweis hierfür hat davon auszugehen, daß dieser Begriff von Kant als ein Begriff konzipiert ist, über den ein Subjekt so verfügt, daß es durch seine Verwendung eine Beschreibung von sich gibt. Ein Subjekt *verfügt* über diesen Begriff, wenn in ihm nicht nur ein Wechsel von Erscheinungen *erfolgt*, sondern wenn es auch ein *Bewußtsein* davon hat, *daß* in ihm ein solcher Wechsel erfolgt. Dieses Bewußtsein kann jedoch durchaus nicht als ein ‚reines' Bewußtsein, also auch nicht als ein Bewußtsein a priori beschrieben werden; *wenn* gilt, daß ein Bewußtsein, das nicht rein ist, kein Bewußtsein a priori sein kann. Zwar *impliziert* das Bewußtsein vom Wechsel der Erscheinungen ein Bewußtsein von der Zeit als der durchgängigen Form, in der Subjekten wie uns sinnliche Anschauungen gegeben werden; es ist jedoch mit dem Bewußtsein von dieser Formbestimmtheit unserer Sinnlichkeit *nicht identisch*. Denn es ist nicht das Bewußtsein von dieser *Form*, sondern vom Auftritt und Abtritt von etwas mit dieser Form nicht Identischem *in* dieser Form. Dieses mit der Form der Zeit nicht Identische ist aber gerade das *empirisch* gegebene Mannigfaltige unserer Sinnlichkeit – dasjenige Mannigfaltige, welches im Unterschied zu dem reinen Mannigfaltigen der Anschauungsform der Zeit selber gerade *nicht* a priori gegeben werden kann und dessen *inhaltliche* Bestimmtheiten und Differenzen auch nicht a priori *antizipiert* werden können. Wenn nun der logische Inhalt des Begriffs des bloßen Wechsels der Erscheinungen unbezüglich auf nur a posteriori gebbare Differenzen der empirischen Anschauung nicht exponibel ist, zeigt dann nicht das Bewußtsein vom Wechsel der Erscheinungen gerade eine Struktur an, die ein Subjekt als die seine bloß *erfährt*, und zwar deswegen, weil es den Wechsel der Erscheinungen nur erfahren, aber nicht a priori antizipieren kann? Dem entsprechend wäre der Begriff des bloßen Wechsels der Erscheinungen ein aus der von Kant ohnehin zugelassenen ‚inneren Erfahrung' abstrahierter *empirischer* Allgemeinbegriff, mithin ein Begriff, mit dem das Subjekt, das über ihn verfügt, das Bewußtsein, der Sachverhalt, der vermittels dieses Begriffs beschrieben wird, könne nur so und nicht anderes bestimmt sein, gerade *nicht* verbindet. Zwar ist es ohne Zweifel so, daß Subjekte wie wir vermittels des Begriffs des bloßen Wechsels der Erscheinungen eine solche Beschreibung von uns geben, der gegenüber wir *de facto* alternativelos sind. Diese Tatsache besagt jedoch durchaus nicht, daß wir uns deshalb schon der *Apriorität* des Bewußtseins vom Wechsel der Erscheinungen bewußt sind.

Gleichwohl muß es möglich sein, dem Begriff des bloßen Wechsels der Erscheinungen unbeschadet der Tatsache, daß sein logischer Inhalt nur mit Bezug auf Differenzen in der empirischen Anschauung exponibel ist, den epistemischen Status eines Begriffs a priori zu sichern. Wäre dies nicht möglich, könnte nicht verstanden werden, daß Kant die 1. Analogie der Erfahrung als ein synthetisches Urteil *a priori* kennzeichnet. Dies bedeutet jedoch, daß man Kants nahezu stereotyp aufgestellte These, nur ‚reine' Anschauungen oder Begriffe seien a priori möglich, als irrige Selbstinterpretation seines eigenen erkenntnistheoretischen Programms preiszugeben hat. Denn der Begriff des bloßen Wechsels der Erscheinungen ist kein ‚reiner' Begriff und muß gleichwohl ein Begriff a priori sein.

Nun muß man sich freilich darüber ins Klare setzen, daß weder für Kant noch der Sache nach ein Argument in Sicht ist, welches dem Begriff des Wechsels der Erscheinungen so etwas wie eine unbedingte Apriorität sichern könnte. Denn weder folgt aus dem Begriff eines Subjekts,

daß ihm überhaupt Vorstellungen als sinnliche *gegeben* werden müssen, noch aus dem Begriff eines Subjekts, dem Vorstellungen als sinnliche gegeben werden, daß ihm diese in der Form der *Zeit* gegeben werden müssen. Soll daher dem Begriff des bloßen Wechsels der Erscheinungen der Status eines empirischen Allgemeinbegriffs, den ein Subjekt von seiner faktischen Verfassung hat, abgesprochen und der Status eines Begriffs a priori zugesprochen werden können, so wird dies nur ‚sub data quadam conditione' geschehen können. Es müßte sich eine Bedingung angeben lassen, unter der ausgeschlossen werden kann, daß ein Subjekt vermittels des Begriffs des bloßen Wechsels der Erscheinungen eine solche Beschreibung von sich gibt, mit der es das Bewußtsein verbindet, es *könnte* über diesen Begriff auch nicht verfügen und dennoch eine Beschreibung von sich geben. Eine derartige Bedingung wäre so beschaffen, daß einem Subjekt unter ihr eine alternative Selbstbeschreibung nicht nur *de facto* nicht möglich, sondern aus *logischen* Gründen unmöglich ist.

Es ist daher zu erwägen, ob nicht in dem Bewußtsein selber, welches durch den Begriff des bloßen Wechsels der Erscheinungen beschrieben wird, eine solche Bedingung aufgewiesen werden kann. Mit ihr müßte eine Struktureigentümlichkeit unserer Subjektivität namhaft gemacht werden, die wir aus logischen Gründen kennen, wenn wir mit diesem Begriff eine Beschreibung von uns geben. Sie wäre eine Bedingung, unter welcher sich der nicht-empirische Charakter des Begriffs vom bloßen Wechsel der Erscheinungen genau dann ergäbe, wenn sich nachweisen ließe, daß (a) kein Bewußtsein, das diese Bedingung erfüllt, ein anderes Bewußtsein als das Bewußtsein vom Wechsel der Erscheinungen sein kann, und (b) diese Bedingung selber nicht mit Sinn eine ‚empirische' genannt werden kann.

Ein solcher Nachweis hat von dem analytisch wahren Satz auszugehen, daß kein Subjekt ein Bewußtsein vom Wechsel seiner empirischen Anschauungen haben kann, ohne ein Bewußtsein von der Zeitlichkeit des Gegebenseins seiner empirischen Anschauungen zu haben. Das Bewußtsein von dieser Zeitlichkeit wäre die gesuchte Bedingung im Sinne des unter (a) formulierten Beweisziels, wenn sich ein zwingendes Argument dafür entwickeln ließe, daß ein Subjekt, welches überhaupt ein Bewußtsein von der *Zeitlichkeit* des Gegebenseins seiner empirischen Anschauungen haben können soll, aus logischen Gründen des Begriffs von diesem Bewußtsein ein Bewußtsein davon haben können muß, daß ihm seine empirischen Anschauungen im *Wechsel* gegeben werden müssen. Es wäre ein Argument für die ‚strenge Allgemeinheit und Notwendigkeit' der Beschreibung, welche ein Subjekt vermittels des Begriffs des bloßen Wechsels der Erscheinungen von sich gibt, *unter der Voraussetzung*, daß es vermittels des Begriffs der Zeitlichkeit des Gegebenseins seiner empirischen Anschauungen eine Beschreibung von sich gibt. Mit ihm wäre der Nachweis erbracht, daß ein Subjekt, das vermittels des Begriffs des bloßen Wechsels der Erscheinungen eine Beschreibung von sich gibt, gegenüber dieser Beschreibung nicht nur de facto, sondern aus logischen Gründen alternativelos ist, sofern es sich überhaupt als ein Subjekt bekannt sein können soll, dem empirische Anschauungen in der Zeit gegeben sind.

Ein derartiges Argument hat die formale Struktur eines ‚transzendentalen Arguments'. Es muß im zu betrachtenden Fall zeigen, daß entwerfbare *Alternativen* zum Bewußtsein vom Wechsel der Erscheinungen, die der Bedingung genügen sollen, Fälle des Bewußtseins von der Zeitlichkeit des Gegebenseins empirischer Anschauungen zu sein, ‚unechte', nämlich in Wahrheit *logisch inkonsistente* Alternativen sind. Ich habe an anderer Stelle drei solche Alternativen vorgestellt und diskutiert, die mir auch den vollständigen Satz solcher Alternativen zu präsentieren scheinen.[2] Diese Diskussion zeigt mit rein analytischen Mitteln, daß die Begriffe vom Bewußtsein von der Zeitlichkeit des Gegebenseins empirischer Vorstellungen und vom Bewußtsein des Wechsels solcher Vorstellungen *logisch äquivalente* Begriffe sind. Von diesem Diskussionsergebnis gehe ich für die Formulierung meiner Schlußüberlegungen aus.

Ein Subjekt, das über den Begriff der Zeitlichkeit des Gegebenseins seiner empirischen Anschauungen verfügt, verfügt mit logischer Notwendigkeit über den Begriff des Wechsels der Erscheinungen (als empirischer Anschauungen). Doch kommt dem Bewußtsein von der Zeitlichkeit des Gegebenseins der empirischen Anschauungen so etwas wie unbedingte oder interne Apriorität nicht zu. Es bezeichnet vielmehr einen bloß subjektiven Sachverhalt, der

auf genau dieselbe Weise ‚etwas ganz Zufälliges' ist, wie dies nach Kant eine uns mögliche Erfahrung von Objekten in Raum und Zeit auch ist. (Vgl. B 765) Gleichwohl muß es möglich sein, dem Begriff von diesem Bewußtsein und damit dem mit ihm äquivalenten Begriff des Wechsels der Erscheinungen den Status eines nicht-empirischen Begriffs zu sichern, um letzterem seine Funktion zu sichern, Subjektbegriff eines synthetischen Urteils a priori zu sein.

Dies kann nicht durch Rekurs auf Kants Lehre von der Zeit als apriorischer Form unserer Anschauung geschehen. Abgesehen von der Anfechtbarkeit dieser Lehre ist dies deshalb unmöglich, weil das Bewußtsein von der Zeitlichkeit des Gegebenseins der empirischen Anschauungen in seinem vollen deskriptiven Bestand nicht das Bewußtsein von dieser Form, sondern von demjenigen ist, was in dieser Form empirisch gegeben ist. Im Begriff des Wechsels der Erscheinungen ist die Zeit des Bewußtseins gerade unter der von dem Formcharakter der Zeit zu unterscheidenden Voraussetzung der Differenz empirischer Daten der Anschauung thematisch. Wenn daher die in Frage stehenden Begriffe als Begriffe a priori rekonstruierbar sein sollen, kann die Rekonstruktion nicht von dieser Differenz als einer bloß *empirisch* gebbaren absehen.

An einer und, soweit ich sehe, nur an dieser Stelle hat Kant selber einen Hinweis darauf gegeben, wie eine solche Rekonstruktion auszusehen hat. Sie ist auch die einzige, an der Kant mit seiner ansonsten stereotypen Forderung nach der ‚Reinheit' von Vorstellungen, deren Begriffe als Termini in synthetischen Urteilen a priori Verwendung finden können sollen, förmlich bricht: „Innere Erfahrung überhaupt und deren Möglichkeit, oder Wahrnehmung überhaupt und deren Verhältnis zu anderer Wahrnehmung, ohne daß irgendein besonderer Unterschied derselben und Bestimmung empirisch gegeben ist, kann nicht als empirische Erkenntnis, sondern muß als Erkenntnis des Empirischen überhaupt angesehen werden, und gehört zur Untersuchung der Möglichkeit einer jeden Erfahrung, welche allerdings transzendental ist." (B 401) – Wenn dies gelten soll, muß es möglich sein, eine Erkenntnis vom eigenen ‚Inneren' zu identifizieren, die von den jeweils *besonderen Unterschieden* in den Gehalten der Wahrnehmungen, in denen sich dies Innere präsentiert, absieht, ohne doch davon abzusehen, daß es sich in *Wahrnehmungen* präsentiert. Man kann solche Wahrnehmungen mit Kant und der Traditionsgeschichte des klassischen Empirismus ‚innere Wahrnehmungen' nennen. Zwar ließe sich im Rekurs auf innere Wahrnehmungen gewiß nicht *begreiflich* machen, daß ein Subjekt Kenntnis von seinem eigenen Inneren hat. Selbstbewußtsein, und sei es auch ‚empirisches Selbstbewußtsein', läßt sich nicht durch Wahrnehmungen erklären. Gibt es aber innere Wahrnehmungen, dann ist das Verhältnis der einen zu der anderen das ihrer zeitlichen Aufeinanderfolge. Genauer gesagt ist dies das grundsätzliche Verhältnis zwischen ihnen, dem gegenüber ein etwaiges Zugleichsein innerer Wahrnehmungen parasitär ist. Denn so, wie nach Kant unsere ‚Apprehension' des Mannigfaltigen der Anschauung ‚jederzeit sukzessiv' ist (vgl. B 225), so ist auch das Bewußtsein davon, daß ein Subjekt in verschiedenen Vorstellungszuständen ist, an das Bewußtsein von einer zeitlichen Folge der inneren Wahrnehmung dieser Vorstellungszustände geknüpft.

Nun kann es gewiß kein Bewußtsein davon geben, daß ein Wechsel der Erscheinungen wirklich erfolgt, ohne daß ‚irgendein besonderer Unterschied' zwischen den Gehalten des innerlich Wahrgenommenen ‚empirisch gegeben ist'. Gleichwohl ist das Bewußtsein vom Wechsel der Erscheinungen einer Beschreibung zugänglich, die es von den einzelnen wirklich erfahrenen Fällen seiner jeweiligen konkreten Instantiierung zu unterscheiden nötigt. Das Argument hierfür geht aus dem Argument dafür hervor, daß ein Subjekt, welches ein Bewußtsein von der Zeitlichkeit des Gegebenseins seiner empirischen Anschauungen haben können soll, ein Bewußtsein davon haben können muß, daß ihm diese Anschauungen im Wechsel gegeben sind. Das Bewußtsein von diesem Wechsel impliziert ein Bewußtsein davon, daß ein Vorstellungszustand, den sich ein Subjekt auf Grund der Präsentation einer bestimmten, d.h. besonderen empirischen Anschauung als den eigenen zuschreibt, weder der *einzige* noch der *letzte* seiner Vorstellungszustände sein *kann*. Denn das Bewußtsein von der bloßen Zeitlichkeit des Gegebenseins der empirischen Anschauungen bezieht sich protensiv auf die Zukunft eines Subjekts als eine noch ausstehende Folge von Zeitpunkten seiner eigenen Zeit,

in der ihm andere Inhalte der empirischen Anschauung als der jetzt gegebene Inhalt gebbar sein müssen. Eine innere Wahrnehmung ist die Wahrnehmung, die sie ist, daher nur unter der Bedingung der in sie selber fallenden Antizipation einer anderen zukünftigen inneren Wahrnehmung, also nur unter der Bedingung der Antizipation eines noch ausstehenden Wechsels der Erscheinungen. Diese Antizipation kann in der Tat nicht so beschrieben werden, daß in ihr selber ein besonderer Unterschied zwischen den empirischen Gehalten der jetzt vorliegenden und der antizipierten inneren Wahrnehmung *gegeben* ist. Denn der besondere Gehalt der empirischen Anschauung, deren innere Wahrnehmung antizipiert wird, ist *ex vi definitionis* von ‚Antizipation' gerade *nicht* gegeben. Was im Bewußtsein von der Notwendigkeit der Möglichkeit eines in Zukunft erfolgenden Wechsels der Erscheinungen antizipiert wird, ist bloß ein Unterschied zwischen empirischen Anschauungen im Modus des Kantischen ‚überhaupt', nur dies, daß der Inhalt einer zukünftigen inneren Wahrnehmung empirisch *anders* bestimmt sein muß als der Inhalt der gegenwärtig gegebenen, – *wie*, bleibt dabei gerade unbestimmt.

Nach dieser Rekonstruktion ist der Begriff des bloßen Wechsels der Erscheinungen primär kein Begriff, mit dessen Hilfe beschrieben wird, was in einem auf zeitliche Weise anschauenden Subjekt wirklich, und d.h. empirisch konstatierbar, vorgeht, sondern der Begriff von einer *Möglichkeit*, die einem solchen Subjekt grundsätzlich garantiert sein muß, wenn es überhaupt ein Bewußtsein von der Zeitlichkeit des Gegebenseins seiner empirischen Anschauungen haben können soll. Das Bewußtsein, welches durch den Begriff des bloßen Wechsels der Erscheinungen beschrieben wird, ist, um es in einem scheinbaren Paradox zu formulieren, eine a priori erfolgende Antizipation genau dessen, dessen jeweils besondere Bestimmtheit gar nicht a priori antizipiert werden kann, weil sie bloß empirisch gegeben werden kann. In diesem, aber auch nur in diesem Sinne ist das Bewußtsein vom Wechsel der Erscheinungen keine innere ‚Erfahrung' von der jeweils faktischen Bestimmtheit, in der sich ein Subjekt auf Grund der nur empirisch gebbaren und bestimmbaren Besonderungen seiner empirischen Anschauungen kennt, sondern eine Bedingung der Möglichkeit jeder derartigen Selbsterfahrung. In diesem, aber auch nur in diesem Sinne ist der Begriff des Wechsels der Erscheinungen kein empirischer Begriff, sondern ein Begriff a priori.

Dies gilt jedoch nur unter der Voraussetzung, daß ein Subjekt ein Bewußtsein von der Zeitlichkeit des Gegebenseins seiner empirischen Anschauung besitzt. *Daß* dies der Fall ist, bleibt auch nach Argumenten für die logische Äquivalenz des Begriffs von diesem Bewußtsein und des Begriffs vom Bewußtsein des Wechsels der Erscheinungen ‚etwas ganz Zufälliges', also ein kontingentes Faktum. Doch ist seine Kontingenz von der Kontingenz der Erfahrungen, die wir wirklich anstellen, noch zu unterscheiden. Denn diese Kontingenz ist nur unter der Bedingung jener möglich. Es gibt keine bestimmten Erfahrungen, seien sie nun Selbst- oder Fremderfahrungen, ohne das ihnen gegenüber invariante Bewußtsein vom Wechsel der Erscheinungen. Wenngleich dies Bewußtsein ebenso kontingent ist wie das Bewußtsein von der Zeitlichkeit der Subjekte, die wir sind, ist es doch mit genausowenig Sinn ein empirisches Bewußtsein zu nennen wie dieses. Es ist im Sinne des oben unter (b) formulierten Beweisziels eine kontingente, aber ihrerseits nicht-empirische Bedingung der Erkenntnis des Empirischen.

Diese Charakterisierung hat jedoch die systematische Pointe, daß dem Bewußtsein vom Wechsel der Erscheinungen so etwas wie eine interne oder unbedingte Apriorität ebensowenig zugeschrieben werden kann wie dem Bewußtsein von der Zeitlichkeit des Subjekts. Doch ist dies auch gar nicht nötig. Es genügt, den Begriff des bloßen Wechsels der Erscheinungen von allen *in specie* so zu nennenden empirischen Begriffen, die wir von uns selber auf Grund jeweils faktisch vollzogener Selbstzuschreibungen besonders gegebener Inhalte der empirischen Anschauung haben, unterscheiden zu können, um ihm die von Kant geforderte *transzendentale* Funktion für eine nicht-empirische Theorie der empirischen Erkenntnis zuzuweisen. Zwar ist er nur der Begriff von der Grundstruktur eines Subjekts, deren kontradiktorisches Gegenteil für dies Subjekt logisch denkbar bleibt – wenngleich nicht als eine auf es selber zutreffende Möglichkeit. Insofern er aber keine Bestimmtheit des Subjekts benennt, die

über die Faktizität seiner eigenen Zeitlichkeit hinaus weitere kontingente Bedingungen seiner Selbsterfahrung in Anspruch nimmt, *kann* er als operativer Begriff einer Theorie fungieren, welche die Bedingungen der einem solchen Subjekt *möglichen* Erfahrung von Objekten (in genauem Unterschied zu dem, was solche Subjekte *de facto* erfahren) rekonstruiert. Der Begriff vom *Wechsel* der Erscheinungen *muß* in einer solchen Theorie aber auch fungieren. Denn er ist genau derjenige Begriff unseres gesamten Begriffsrepertoires, der im Unterschied zu Begriffen von einer ‚reinen' Anschauung und zu Begriffen von einer kategorialen Synthesis a priori auf das bloß empirisch gebbare Mannigfaltige und dessen Differenzen im Modus des ‚überhaupt' kraft seines logischen Inhalts bezogen ist. Er enthält eben daher die formalen subjektiven Bedingungen der Anwendung von Kategorien der *Relation* auf Gegebenheiten unserer Sinnlichkeit. Denn Kategorien der Relation lassen sich auf unsere Sinnlichkeit nur dann beziehen, wenn sie als Begriffe von der notwendigen Verbindung dessen aufgefaßt werden, was uns als eine Mannigfaltigkeit *empirisch* gegebener Daten in der Form der Zeit präsentiert wird. An den Anschauungsformen von Zeit und Raum als solchen haben Relationskategorien keinen Anhalt. Sie gehen nicht auf Zeit und Raum als Gegenstände vorgestellt, sondern auf etwas in Zeit und Raum als Gegenstand vorgestellt: auf das ‚Dasein' von Objekten *in* Raum und Zeit. Das ist der systematische Grund dafür, daß der Begriff des bloßen Wechsels der Erscheinungen nicht nur als Subjektbegriff der 1. Analogie der Erfahrung (und damit als der Grundbegriff aller Analogien der Erfahrung) fungieren kann, sondern auch fungieren muß. Diese Notwendigkeit kann jedoch nur dann einsichtig gemacht werden, wenn jene Möglichkeit verstanden werden kann. Die voranstehend entwickelten Überlegungen wollen als ein Versuch aufgefaßt werden, einem solchen Verständnis näherzukommen. Sie gelten der Frage, auf welche Weise dem Begriff des Wechsels der Erscheinungen die von Kant geforderte Apriorität gesichert werden kann, ohne die für seinen logischen Inhalt konstitutive Beziehung auf Differenzen der empirischen Anschauung aus ihm zu exportieren. Wo dies geschieht, begibt sich der Interpret jeder Möglichkeit, die transzendentale Funktion des Systems der Analogien der Erfahrung im Aufbau der Kantischen Epistemologie und die von Kant für sie angebotenen Beweisgründe zu verstehen.

ANMERKUNGEN

[1] Die avancierteste Rekonstruktion von Teilen der transzendentalen Deduktion hat vorgelegt D. Henrich, *Identität und Objektivität*. (Heidelberg 1976).

[2] Cramer, K., *Nicht-reine synthetische Urteile a priori. Ein Problem der Transzendentalphilosophie Immanuel Kants* (Heidelberg 1985), Kap. 10: Formale Probleme des Begriffs des bloßen Wechsels der Erscheinungen.

* * *

DAS LEIB-SEELE-PROBLEM, EPISTEMOLOGISCH GESEHEN

François Bonsack
Universität Neuchâtel
Institut für Methodologie, Biel

Die Unterscheidung zwischen Seele und Leib kann auf verschiedenen Ebenen gemacht werden:

a) auf der ontologischen Ebene, die dann auf die metaphysisch-religiöse erweitert werden kann;

b) auf der ethisch-moralischen Ebene;

c) auf der epistemologischen Ebene.

a) *Ontologische Ebene*

Der ontologische Dualismus postuliert zwei Substanzen, z.B. Descartes „res extensa" und „res cogitans", die beide objektiv existieren und sich durch verschiedene Eigenschaften unterscheiden, wie Eisen vom Sauerstoff, ein Proton von einem Photon oder ein Teilchen von einem Feld.

Die ontologische Dualität von Seele und Leib hatte den großen Vorteil, daß sie eine naheliegende Interpretation des Todes bot. Daher kam sie in den verschiedensten Kulturen vor.

Was unterscheidet nämlich einen lebenden Körper von einem Leichnam? – Materiell scheinbar nichts: die Organe, die Gewebe, die Zellen sind die gleichen, das Blut ist immer da (aber es zirkuliert nicht mehr), auch die Wärme ist zunächst noch da, kurz alles Leibliche besteht und doch funktioniert nichts mehr. Beweglichkeit und Empfindlichkeit sind verschwunden, der Körper ist nicht mehr fähig, seine Wärme aufrechtzuerhalten, und er zerfällt dann mit der Zeit. Es fehlt ihm offenbar etwas, das er besaß solange er am Leben war, obschon er materiell noch alles zu haben scheint. Was ihm fehlt, ist also notwendigerweise unkörperlich, es ist die Seele, die ihn funktionieren ließ, ihn empfindlich und beweglich machte und ihn beherrschte. Eine einheitliche Seele macht den Leib lebendig und löst sich von diesem im Augenblick des Todes.

Dies war der erste Schritt. Aber sogleich kam der zweite: Was ist diese Seele, die sich vom Körper löst? – Wenn sie das ist, was den Leib beweglich und empfindlich macht, liegt es nahe, sie mit dem bewußten Ich zu identifizieren, das eben auch das ist, was dem Leib Befehle gibt und von ihm Empfindungen bekommt. Die Seele wurde also auf das bewußte Ich eingeschränkt, ungeachtet dessen, daß man dabei einen großen Teil der unbewußten Funktionen fallen ließ, die beim Tode verschwinden und z.B. im Koma oder in der Narkose noch vorhanden sind.

Was geschieht weiter mit der abgelösten Seele? – Die Antwort auf diese Frage geben Religion und Metaphysik: sie könnte sich verflüchtigen, sie könnte erhalten bleiben, eventuell in andere Körper eintreten (von Menschen oder von Tieren), nach anderen Regionen wandern usw. Der Erhaltungstrieb, der „Wille zur Ewigkeit", wie Nietzsche sagt, und auch der Wunsch, geliebte Verstorbene wieder zu treffen, bevorzugen natürlich Lehren, welche die Unsterblichkeit der Seele verkünden.

Man muß aber bezweifeln, ob eine solche Begründung für die Unterscheidung von Leib und Seele heute noch bestehen kann. Die Biologie hat uns zu einer ganz anderen Interpreta-

tion des Todes geführt. Dem „Leben“ wurden große Funktionen unterstellt, die alle notwendig sind für dessen Erhaltung: Verdauung, Atmung, Stoffwechsel, Blutzirkulation, Nervenübertragung, Empfindlichkeit, Muskelkontraktion usw., und jede dieser Funktionen kann in unzählige kleinere zerteilt werden, die alle an bestimmte Gewebe, Zellen, chemische Stoffe gebunden sind. Das Leben einer Zelle ist eine lokale Angelegenheit, die auf ein günstiges physikalisch-chemisches Milieu angewiesen ist, welches seinerseits durch das Zusammenwirken der meisten Organe erhalten wird. Der Tod folgt auf das Versagen einer oder mehrerer der großen Funktionen, welche zur Bewahrung dieses Milieus nötig sind.

Dieses Bild, das uns die heutige Biologie vom Tode entwirft, ist sehr weit von der Trennung einer körperlichen und einer unkörperlichen Substanz entfernt: die einzelnen Zellen sterben nach und nach ab, wenn die physikalisch-chemischen Bedingungen ihres Weiterlebens nicht mehr vorhanden sind, so daß der Genfer Zoologe Emile Guyénot eine seiner Vorlesungen „Teilweises Überleben nach allgemeinem Tod“ betiteln konnte.

b) *Ethisch-moralische Ebene*

Wenn man Plato's „Phaidon“ liest, wird einem sehr bald klar, daß die Seele der Vernunft gleichgestellt wird und daß nicht nur die „niederen“ Triebe, sondern auch Affekte und Leidenschaften dem tierischen Leib zugeschrieben werden. Das ist auch verständlich, sobald man die Seele mit dem bewußten Ich identifiziert hat: dieses befiehlt nicht nur den Muskeln, sondern auch den leiblichen Begehren. Und im Christentum, z.B. im Römerbrief, kommt noch eine moralische Wertung hinzu: das Fleisch wird mit der Sünde und der Geist mit dem Heil gleichgestellt.

Die axiologisch-moralische Leib-Seele-Unterscheidung ist an sich logisch ganz unabhängig von der ontologischen Leib-Seele- oder allgemeiner Geist-Materie-Unterscheidung: man kann dem Geist und der Kultur einen hohen Wert beimessen, auch wenn man in ihnen nicht eine besondere, von der Materie verschiedene Substanz sieht, sondern eine Funktion, welche durch eine besondere Anordnung der Materie möglich gemacht wird – sie sind deshalb nicht weniger wertvoll. Und aus der Vernunft eine höhere biologische Funktion machen hindert diese nicht daran, Leidenschaften, Triebe und Körper zu beherrschen. Natürlich muß man dann auf eine transzendente Begründung der Moral (z.B. mit Hilfe eines letzten Gerichtes) zugunsten einer immanenten verzichten, was ihr beim heutigen Zeitgeist nur zugute kommen kann.

c) *Epistemologische Ebene*

Ich komme nun zu meinem eigentlichen Thema: die Leib-Seele-Unterscheidung auf der epistemologischen Ebene.

Als Ausgangspunkt meiner Ausführungen nehme ich folgende einleuchtende Feststellung: Wir haben keinen unmittelbaren Zugang zur Außenwelt; alles, was wir über sie erfahren, erreicht uns *mittelbar* über unsere Empfindungen, Wahrnehmungen, Vorstellungen. Es ist uns nicht möglich, unsere innere Wahrnehmung einer Kirsche mit einer an sich seienden, außenstehenden Kirsche zu vergleichen und so zu prüfen, ob diese Wahrnehmung die Kirsche richtig wiedergibt. Denn die außenstehende Kirsche ist uns eben nur über unsere Wahrnehmung zugänglich. Wir können nur Wahrnehmungen mit Wahrnehmungen vergleichen (z.B. jetzige mit früheren oder mit späteren, gesehene mit gehörten), Wahrnehmungen mit Begriffen, mit Aussagen anderer über ihre eigenen Wahrnehmungen, mit Wahrnehmungen von Bildern oder von Meßinstrumenten, welche besondere Eigenschaften der Kirsche bestimmen würden. Niemals aber die Wahrnehmung eines Objekts mit dem Objekt selbst. Das Objekt steht uns nicht zur Verfügung als Maßstab dessen Wahrnehmung.

Daher kann die Richtigkeit einer Wahrnehmung nicht an deren Übereinstimmung mit dem

Objekt geprüft werden, sondern nur in einem komplexen Vergleich zwischen verschiedenen Empfindungen, bekannten Begriffen und Gesetzen und einer Abwägung der verschiedenen Möglichkeiten in der gegebenen Lage.

Diese Feststellung kann so interpretiert werden, daß uns nur ein Bild in uns von der Außenwelt bekannt ist, und daß die Existenz dieser Außenwelt bezweifelt werden kann, dies umso mehr, als uns die Sinne manchmal täuschen und wir auch träumen, d.h. innere Bilder sozusagen wahrnehmen, denen in der Außenwelt nichts entspricht.

So schreibt Descartes in seinen *Prinzipien der Philosophie* (Teil I, § 8):

> Examinant ce que nous sommes, nous qui pensons maintenant qu'il n'y a rien endehors de notre pensée qui soit véritablement ou qui existe, nous connaissons manifestement que, pour être, nous n'avons pas besoin d'extension, de figure, d'être en aucun lieu, ni d'aucune autre telle chose que l'on peut attribuer au corps et que nous sommes par cela seul que nous pensons; et que par conséquent la notion que nous avons de notre âme ou de notre pensée précède celle que nous avons du corps, et qu'elle est plus certaine, vu que nous doutons encore qu'il y ait aucun corps au monde, et que nous savons certainement que nous pensons.
> (Denn wenn man prüft, wer wir sind, wir, die wir jetzt davon überzeugt sind, daß es nichts außerhalb unseres Bewußtseins gibt, das wahrhaft ist oder existiert, so sehen wir deutlich, daß wir weder Ausdehnung, noch Gestalt, noch Ort, noch irgend etwas anderes, das man dem Körper zuschreiben kann, brauchen, um zu sein, und daß wir allein dadurch sind, daß wir denken. Folglich geht der Begriff, den wir von unserer Seele oder von unserem Denken haben, demjenigen voran, den wir vom Leib haben, und er ist gewisser, da wir ja noch bezweifeln, daß es auch nur einen einzigen Körper in der Welt gebe, aber sicher wissen, daß wir denken.)

Dies führt uns zu einer epistemologischen Unterscheidung von Seele und Leib: Seele ist das, was uns unmittelbar sozusagen von innen zugänglich ist; Außenwelt und der zu ihr gehörende Leib ist das, auf dessen Existenz erst aus diesem unmittelbar Gegebenen mittelbar geschlossen wird. Unsere sichere Erkenntnis erstreckt sich nur auf unsere eigene Seele; zwischen ihr und der Welt liegt eine unüberbrückbare epistemologische Kluft.

Die Hauptfrage ist nun: entspricht dieser epistemologischen Unterscheidung eine ontologische? Kann man aus der Tatsache schließen, daß uns die Seele von innen erreichbar ist und der Leib sozusagen von außen, daß die Seele aus einem anderen Stoff aufgebaut ist als der Leib?

Nichts zwingt uns dazu. Und dieser Schluß hätte nicht geringe Nachteile, denn es würde die epistemologische Kluft in eine ontologische verwandeln; man würde dann nicht mehr verstehen, wie so verschiedenartige Substanzen aufeinander wirken können. Das bereitete schon Descartes und seinen Nachfolgern große Schwierigkeiten, welche diese zum Teil sehr unbefriedigende Lösungen vorschlugen.

„Seele" und „Leib" im Sinne Descartes müssen nämlich aufeinander wirken können, einerseits um die kausale Verbindung von objektiven Ereignissen auf unsere innere Sphäre zu erklären, und dies nicht nur bei den Sinnesempfindungen, sondern auch beim Einfluß von Drogen auf unseren Bewußtseinszustand. Andererseits muß man der Wirksamkeit unseres Willens auf die Außenwelt Rechnung tragen. Irgendwie müssen also Seele und Leib, oder Geist und Materie auf einer gemeinsamen ontologischen Ebene gelegen sein, damit ihre Wechselwirkungen interpretiert werden können.

Das heißt, daß das, was ich hier sehr vage mit „Seele" bezeichnet habe – nämlich die subjektive Sphäre, alles, was von innen her zugänglich ist – irgendwie reifiziert, objektiviert, als zur Welt gehörend betrachtet werden muß. Dies ist unumgänglich, wenn man daran festhält, daß es nur auf der ontologischen, objektiven Ebene echte Kausalbeziehungen gibt und daß es kontinuierliche Kausalketten gibt die von der Welt zu unserer subjektiven Sphäre und umgekehrt von letzterer zur Welt führen.

Aber wirft nicht eben diese Behauptung gewaltige Probleme auf? Man kann möglicherweise einig werden, daß es so sein muß, aber es bleibt die große Aufgabe zu erklären, *wieso* es möglich ist.

Hier einige Bemerkungen zu diesen Problemen:

a) Das *Objekt* unserer Wahrnehmung darf nicht mit dem *Mittel*, mit dessen Hilfe wir wahr-

nehmen, verwechselt werden. Daß ich einen Baum über das Bild wahrnehme, das auf meine Netzhaut abgebildet wird, heißt nicht, daß ich nichts anderes als dieses Bild sehe, daß es das *Objekt* meines Sehens ist – welche Behauptung nur die unsinnige Interpretation zuließe, daß dort, wo ich vermeintlich einen Baum sehe, in Tat und Wahrheit nur ein Bild des Baumes steht. Das Bild auf meiner Netzhaut ist nur ein Glied in der Kette, die zur Wahrnehmung führt; Objekt der Wahrnehmung ist nicht ein Zwischenglied, nicht einmal das Endglied dieser Kette, sondern deren „Anfangs"-glied.

Das Objekt meines Sehens ist also der Baum, nicht meine Wahrnehmung vom Baum. Man sollte nicht die heimliche Vorstellung hegen, daß ein zweites Subjekt in unserem Kopf steckt, welches das beobachtet, was uns vom äußeren Objekt durch unsere Sinne und Nerven übermittelt wird, nämlich bloß Nervenreize verschiedener Form auf verschiedenen Leitungen. Vielleicht tönt das naiv, doch bin ich nicht sicher, ob die eine oder andere Erkenntnislehre nicht in einer solchen Vorstellung ihre Wurzel hat, und es gibt Physiologen, die sich ernsthaft mit der Frage befaßt haben, warum uns die Objekte aufrecht erscheinen, obgleich sie auf der Netzhaut umgekehrt abgebildet sind.

b) Müssen zwei Substanzen wesensgleich sein, um aufeinander wirken zu können? Schließt die Hypothese, daß Seele und Leib wesensfremde Substanzen sind, jede Wechselwirkung zwischen ihnen aus? Sind nicht z.B. ein geladenes Teilchen und ein elektrisches Feld sehr verschiedene Entitäten, was aber das Teilchen nicht daran hindert, ein Feld zu erzeugen und dieses seinerseits, auf ein Teilchen eine Kraft auszuüben?

Die Antwort auf diese Frage hängt vom Status ab, das man der Kausalität gewährt. Gewöhnlich verbindet diese objektive Ereignisse, sie ist sozusagen physikalischer Natur. Ob es eine andere Art der Kausalität gibt als diese physikalische, ist nicht erwiesen. Wenn sie aber materielle und immaterielle Ereignisse verbinden soll, muß sie halb materieller, halb immaterieller Natur sein, was nicht leicht zu lösende metaphysische Probleme stellt.

Etwas sollte man aber m. E. mindestens festhalten, nämlich daß die Kausalität etwas Objektives ist, und daß sie Elemente objektiver Natur miteinander verbindet: z.B. sind Teilchen und Feld beide objektive Entitäten.

c) Niemand sollte daran Anstoß nehmen, daß eine und dieselbe Entität, zur ontologischen Ebene gehörend, epistemologisch auf ganz verschiedene Arten erfaßt werden kann. Es ist sogar die Regel. Ein und derselbe Kreis kann uns als verschieden abgeplattete Ellipsen oder sogar als ein Geradensegment erscheinen. Ein und derselbe Würfel kann gesehen und betastet, ein Blitz gesehen und gehört werden.

Nichts hindert also die Seele, einerseits als ein subjektiv Empfundenes zu erscheinen und andererseits eine wie der Leib zur Welt gehörende objektive Realität zu besitzen. Das scheint besonders natürlich für die subjektive Sphäre anderer, die ich nicht von innen empfinden kann wie meine eigene, sondern auf die ich nur über deren objektives Verhalten schließen kann. Epistemologisch hat mein Bewußtsein nicht nur einen gewaltigen Vorrang, es breitet sich sogar über die ganze Szene aus und duldet nichts außer sich; ontologisch ist es aber nur eines unter anderen und hat da ein objektives Sein wie die anderen belebten und unbelebten Körper.

d) Wieso darf ich von einer ontologischen Ebene sprechen, wenn ich doch von der Feststellung ausgegangen bin, daß uns die Welt nur über Wahrnehmungen, d.h. über private Ereignisse, erreichbar ist?

Dieses Vorgehen enthält aber gar keinen Widerspruch, wenn man mit Einstein die These vertritt, daß die ontologische Ebene nicht etwas primär Gegebenes ist, sondern etwas, das ich, von den Wahrnehmungen ausgehend, konstruiere, als Erklärung eben dieser Wahrnehmungen. *Im* Modell sind die ontologischen Entitäten Ursache der modellierten Wahrnehmungen, aber *das* Modell habe ich nach den ursprünglichen Wahrnehmungen konstruiert. Auf der Metaebene sind Wahrnehmungen epistemologisch das erste und Ontologie das letzte, im Modell sind die ontologischen Entitäten das erste und verursachen kausal die Wahrnehmungen.

Die ontologische Ebene unterscheidet sich von der objektiven dadurch, daß letztere noch

epistemologischer Natur ist, d.h. je nach individuellem Subjekt und dessen Wissensstand variieren kann, die ontologische Ebene aber als etwas postuliert wird, das sich nicht mit dem Aneignen neuen Wissens ändern würde und vom einzelnen Erkennenden nicht abhängig wäre. Ontologie ist also nur ein praktisch unerreichbarer Grenzbegriff, der jenseits aller heutigen und künftigen objektiven Wissen der einzelnen Subjekte liegt. Wir postulieren sie, sie ist im Prinzip nicht unerkennbar, aber uns steht nur unser individuelles objektives Wissen zur Verfügung, das ihr gegenüber nur lückenhaft und teilweise sogar falsch sein kann. Noch mehr: dort, wo unser persönliches Modell gegebenfalls mit ihr übereinstimmen würde, wüßten wir es nicht einmal. Von keinem Teil unseres Wissens können wir behaupten, er entspreche endgültig der Ontologie. Das einzige, was wir wissen ist, daß diese sich nicht ändert, wenn bloß unser Wissen erweitert oder verbessert wird.

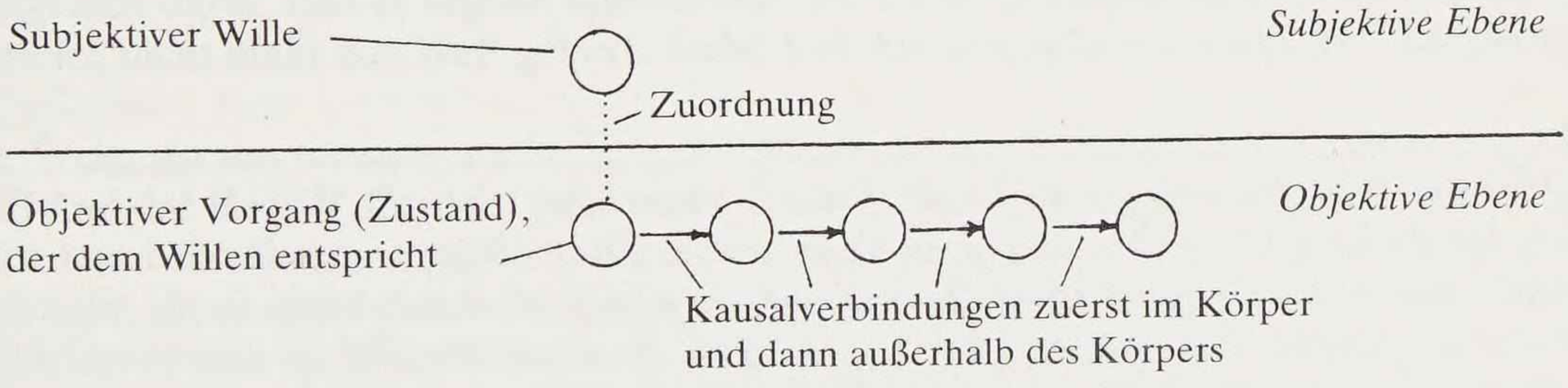
e) Objektivieren bedeutet nicht „wissenschaftlich erklären". Man kann einem Ereignis eine objektive Realität zuschreiben, auch wenn man es nicht im Einzelnen erklären und auf bekannte Gesetze zurückführen kann.

Ähnlich braucht man, um eine kausale Beziehung zwischen zwei Ereignissen aufzuweisen, nicht alle Glieder der Kette erklärt zu haben. Es gibt eine empirisch-pragmatische Kausalität: ich beweise, daß B Wirkung von A ist, wenn ich B nach Belieben erzeugen oder hindern kann, indem ich auf A wirke. Ich sage, daß meine subjektive Wahrnehmung eines Objekts von diesem Objekt verursacht wird, weil ich meine Wahrnehmung erzeugen oder aufheben kann, indem ich es anschaue oder mich von ihm abwende; ich sage, daß mein subjektiver Schmerz vom Stich dieser Nadel verursacht wird, weil ich es jedesmal empfinde, wenn die Nadelspitze in meine Haut eindringt.

Die Erklärung einer kausalen Beziehung z.B. mit Hilfe eines wissenschaftlichen Gesetzes bringt zu diesem empirisch-pragmatischen Begriff der Kausalität nichts hinzu, denn die Aufstellung des Gesetzes ist nichts mehr als die Legalisierung einer solchen vorher konstatierten kausalen Verbindung.

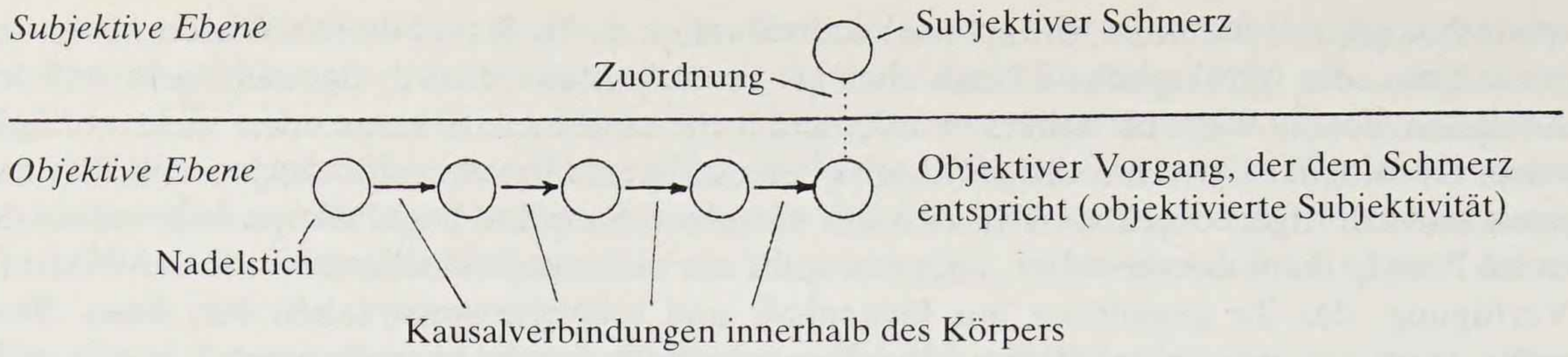
f) Die Zuordnung von Elementen zweier Ebenen ist streng von einer kausalen Beziehung zwischen diesen einander zugeordneten Elementen zu unterscheiden.

Gewissen subjektiven Ereignissen (z.B. Empfinden eines Schmerzes) werden objektive Ereignisse (Verletzung der Haut, Nervenübertragung, Vorgänge im Zentralnervensystem) zugeordnet. Das heißt aber nicht, daß der neurophysiologische Vorgang in meinem Zentralnervensystem meinen subjektiven Schmerz *verursacht*. Verursacht wird etwas auf der *ontologischen* Ebene, dem epistemologisch einerseits ein subjektiver Schmerz, andererseits ein neurophysiologischer Vorgang zugeordnet ist.



Subjektive und objektive Ebenen sind komplementäre *epistemologische* Ebenen.
Subjektiver und „objektiver Schmerz" sind getrennt auf der epistemologischen Ebene, aber verschmelzen zu einer einzigen Entität auf der ontologischen Ebene.

Man könnte für den Willensakt ein ähnliches Schema zeichnen:



Aber der empirisch-pragmatische Begriff der Kausalität braucht nicht an der Grenze der Objektivität Halt zu machen: es können die objektive Kausalkette und die Zuordnung zwischen objektiver und subjektiver Ebene auf einmal erfaßt werden: die Tatsache, daß ich einen subjektiven Schmerz beliebig erzeugen und beheben kann, indem ich mich steche oder aufhöre, mich zu stechen, beweist die Kausalkette auf der ontologischen Ebene und die Zuordnung eines subjektiven Schmerzes zu diesem Ereignis.

Zusammenfassend würde ich sagen, daß die Leib-Seele- und die Geist-Materie-Unterscheidung, welche die erste erweitert, verwischt wurden durch eine Verschmelzung aller Ebenen: einerseits wurde eine vorwissenschaftliche ursprüngliche Erklärung zu einer metaphysisch-religiösen, zusätzlich belastet mit moralischen Wertungen, andererseits hat man von einer epistemologischen Differenz auf eine ontologische geschlossen. Eine saubere Trennung all dieser Gesichtspunkte kann zu einem klareren Verständnis der Probleme führen oder sogar manche von diesen überhaupt verschwinden lassen.

* * *

DAS METAPHYSISCHE SUBJEKT: DIE STUFE EINER LEITER[1]

Paul Hübscher
Universität Bern

Nur Logisches ist logisch faßbar. Diese Tautologie ist das Fundament von Wittgensteins Erkenntnistheorie im *TLP*. Sie erklärt auch die Stellung des Subjekts in Wittgensteins *TLP*. Das Subjekt ist das logische Ordnende. Doch in der Logik gibt es kein Subjekt, sondern nur Objekte und ihre Relationen (vgl. *TLP* 5.631). D.h., wenn das Subjekt die Objekte logisch ordnet, kommt es zum Schluß, daß es selbst nicht existiert. Das erkenntnistheoretische Subjekt schafft sich selber ab. Es ist sich selber ja nur *logisch* faßbar, als Ding für es, nicht als Ding an sich, als Subjekt. Die Erkenntnistheorie des *TLP* akzeptiert nur Dinge-für-das-Subjekt, aber kein Subjekt. Doch gerade indem sie das Subjekt eliminiert, kreiiert sie es. Dem Begriff ‚Subjekt' in der Bezeichnung ‚Dinge-für-das-Subjekt' muß logischerweise etwas entsprechen, sonst wäre er nicht im Sprachbild. Die Existenz dieses Subjekts kann zwar noch gesehen werden, aber sie ist nicht mehr logisch faßbar. Wir stoßen im *TLP* zweimal auf eine Grenze des Erkennens: bei den Dingen (*TLP* 6.44) und beim Subjekt. Das macht den *TLP* über weite Strecken zu einer metaphysischen Abhandlung.

Auch Logik ist in gewissem Sinn metaphysisch. Sie definiert die Welt, ist aber nicht materieller Teil davon. Sie ist metaphysisch im Sinne von Kants ‚metaphysica generalis': eine Aussage über die Gegenstände möglicher Erfahrung. Doch der *TLP* stößt in ein Gebiet vor, das zur ‚metaphysica specialis' gehört, die als Wissenschaft laut Kant nicht möglich sein soll. Hier geht es letztlich um das höchste Wesen – Gott.

Die Schwierigkeit bei der Behandlung der Metaphysik ist, daß die Sprache als logisches Bild der Welt gegenüber Metaphysischem versagt. Die Grenze, die der *TLP* zwischen Welt und ‚metaphysica specialis' zieht, ist wesentlich eine sprachliche Grenze. Das erkennende Subjekt kann die Grenze transzendieren, aber es wird die Sprache zurücklassen müssen. Der *TLP* versucht, die Grenze aufzuzeigen, indem er dem Subjekt nachweist, was es hat und was es nicht hat. Dieses an den Jahresabschluß eines Kaufmannes erinnernde Vorgehen zeitigt folgendes Resultat: Das Subjekt hat die Welt nur, insofern sie logisch geordnet ist. Wo die Logik aufhört, hört die Welt-für-das-Subjekt auf. Die verblüffende Einsicht des *TLP* ist, daß die Logik aufhört, wo sie beginnt – im die Welt ordnenden Subjekt. Daß das Subjekt die Grenze ist, *zeigt* sich darin, daß es logisch nicht erfaßt ist. Es ist „metaphysisch" (*TLP* 5.633), da es als Grenze nicht mehr zur Welt gehört. Gehört es zur ‚metaphysica specialis'? Ich behaupte: Nein.

Die Welt, die das metaphysische Subjekt ordnet, ist der Mikrokosmos (*TLP* 5.63). Dieser Begriff und der Begriff ‚Grenze' (und nicht ‚Ende'!) der Welt suggerieren, daß es noch etwas jenseits des Mikrokosmos geben muß: einen Makrokosmos. Dieser Makrokosmos muß alogisch sein, da er sonst durch die Logik in den Mikrokosmos eingeordnet würde. Dennoch manifestiert er sich im Mikrokosmos. Es zeigt sich, daß nicht alles dort logisch und erklärbar ist. Zum Beispiel sind Naturgesetze bloße Beschreibungsnetze. Daß sich ein Gegenstand nach ihnen richtet, ist sozusagen reine Gefälligkeit von ihm. Das Subjekt erfährt, daß seine Welt eine Grenze gegen außen hat: „Die Welt ist unabhängig von meinem Willen." (*TLP* 6.373) Diese Welt der unabhängigen Objekte ist der Makrokosmos. Wittgenstein nennt ihn nicht, da er darauf beharrt, daß über ‚metaphysica specialis' nicht gesprochen werden kann. Er weist nur auf, wo sich der Makrokosmos im Mikrokosmos zeigt. Hat das ordnende Subjekt erfaßt, daß es etwas von ihm Unabhängiges gibt, so hat es sich als „metaphysisches Subjekt" (*TLP* 5.633) erfaßt, d.h. es begreift seinen Mikrokosmos als eingegrenztes Ganzes. Zugleich

begreift es sich selber als Grenze. Dies ist die „Anschauung der Welt sub specie aeterni", das „Gefühl der Welt" (*TLP* 6.45). Das ‚Gefühl der Welt' ist ein nicht-logisches Erfassen der Welt. Es ist Erkenntnis ohne Kategorien. Hier ist das Subjekt jenseits des Mikrokosmos und damit jenseits der Sprache.

Es ist seine an Frege orientierte Abbildtheorie der Bedeutung, die es Wittgenstein erlaubt, diese Leiter vom Mikrokosmos in den Makrokosmos zu bauen. Für ihn wie für Frege ist die Bedeutung eines Zeichens an real Gegebenes gebunden. Doch der Sinn eines Zeichens kann außerhalb des real Gegebenen liegen. Wittgenstein erblickt hier die Möglichkeit, die ‚metaphysica specialis' in der Abbildtheorie des *TLP* einzuordnen. Ein Zeichen kann in seinem *Sinn* auf die ‚metaphysica specialis' hinweisen. Es ist allerdings eine besondere Art von Zeichen. Wittgenstein legt Wert darauf, daß das Satzzeichen eine Tatsache ist und daß Tatsachen als Zeichen aufgefaßt werden können, ja daß nur Tatsachen einen Sinn haben können (vgl. *TLP* 3.14–3.1431). Einen auf die ‚metaphysica specialis' hinweisenden Sinn kann nun nur eine ‚Tatsache großen Ausmaßes' haben: das Leben. Wittgenstein identifiziert damit den logischen, von Frege inspirierten Gebrauch von ‚Sinn' mit dem der Alltagssprache, wo ‚Sinn' auch ‚Zweck' oder ‚Ziel' konnotieren kann.

Damit haben wir eine Stufenfolge von Subjekten: Das Leben wird von einem empirischen Subjekt geführt. Dieses wird im *TLP* abgebildet als das sprachliche Subjekt ‚Ich'. Die Abbildung führt aber als Verallgemeinerung auf eine höhere Stufe. ‚Ich' bezeichnet nun das erkenntnistheoretische, logische Subjekt. (Das logische Subjekt ist sozusagen eine Variable, die den Wert eines empirischen Subjektes annehmen kann). Dieses logische, erkenntnistheoretische Subjekt regiert ordnend seine Welt, den Mikrokosmos. Wittgenstein nennt es das metaphysische Subjekt. Es ist im selben Sinne metaphysisch, wie die Logik metaphysisch ist. Zwar nicht Teil des Mikrokosmos, aber mit ihm verbunden, sind sie der ‚metaphysica generalis' zuzuordnen. Hier endet das Welt-Bild des *TLP*, da hier für ihn die Sprache endet. Doch er hinterläßt Pfeile, die über das Ausgesprochene hinausdeuten. Der Begriff ‚Mikrokosmos' weist auf den Begriff ‚Makrokosmos': wenn das Bild, das die erkennend geordnete Welt gibt, richtig ist, muß es etwas geben, das sich nicht erkennend ordnen läßt. Es gibt aber für den *TLP* keine Welt ohne Ordnung. Somit können wir folgende Schlüsse ziehen: Auch der prima facie nicht geordnete Makrokosmos verfügt über Ordnung. Er verfügt damit über ein ordnendes Subjekt. Wie das erkenntnistheoretische Subjekt gegenüber dem empirischen auf einer höheren Stufe steht, steht das Subjekt des Makrokosmos gegenüber dem erkenntnistheoretischen höher. Der Makrokosmos ist somit das Gebiet der ‚metaphysica specialis', sein Subjekt ist Gott.

Sprachlich ist Gott nicht zugänglich. Die Philosophie und insbesondere der *TLP* ist der Versuch, dieses Nichtsprachliche aufzuzeigen, indem das Sagbare vom Unsagbaren unterschieden wird. Das Sagbare aber sind Sätze der Naturwissenschaft, nur diese haben Bedeutung. Andere Sätze gibt es nicht, und damit keine Philosphie. Die einzige Art, Philosophie zu lehren, ist, den Kandidaten darauf hinzuweisen, daß er entweder Naturwissenschaft treibt oder bedeutungslose Sätze äußert (*TLP* 6.53). ‚Metaphysica specialis' wird nicht mit unbedeutenden sprachlichen Zeichen erfaßt, sondern nur mit dem großen Symbol eines ganzen Lebens.

Diese Einsicht hat ihre Konsequenzen auf die Selbsteinschätzung des *TLP*. ‚Logik' und ‚Philosophie' sind nicht nur im Titel sondern in der ganzen Abhandlung prominent. Doch die Logik entpuppt sich als in ihren Gründen nicht begründbar (Wittgenstein nennt kein Beispiel eines Elementarsatzes oder eines Namens). Die Philosophie existiert nur als Methode, sich selber zu überwinden. Dieses Resultat ist paradox, unsinnig (*TLP* 6.54), nonsense. Es ist nicht sinnlos. Der *TLP* hat im Gegenteil sehr viel Sinn, weil er zeigt, in welchem Verhältnis das empirische Subjekt zu Gott steht. Indem er den wesentlich abbildenden Charakter der Zeichen aufzeigt, wird er zu einer Leiter. Das empirische Subjekt erkennt sich und seine Grenzen in seinem Abbild, dem erkenntnistheoretischen („metaphysischen") Subjekt. Es erkennt auch, daß es ein makrokosmisches Subjekt geben muß. Paradox und deshalb unsinnig ist der *TLP*, weil das Subjekt am Schluß erkennt, daß es der Leiter gar nicht bedurft hätte, daß ihm

die Logik des *TLP* bei der Gewinnung des Sinns des Lebens nicht helfen *kann*. Die ‚metaphysica specialis' ist nur durch ein Heraustreten aus der Logik zu begreifen. Dieses Heraustreten geschieht in der mystischen Anschauung der Welt als-begrenztes-Ganzes. Doch nun kann nichts mehr gesagt werden, da mit der Logik auch die Sprache verschwindet.

ANMERKUNG

[1] Es handelt sich hier um eine Kurzfassung der ausführlichen Diskussion desselben Themas in: P. Hübscher, *Der Einfluß von Johann Wolfgang Goethe und Paul Ernst auf Ludwig Wittgenstein* (Bern 1985).

* * *

ON THE AGREEMENT OF HUMAN MIND AND THE REASON OF NATURE

Adolf Hübner
Austrian Ludwig Wittgenstein Society, Kirchberg am Wechsel

I

We suggest accepting as objectively correct (true)
a) that within the existing universe (cosmos, physical reality, nature, "world"), which has the *property* "to be in motion", nothing can exist in a state of (absolute) rest. ("Nothing can be static"),
b) that, because of the property of the existing universe "to be in motion" a class of objects "being at (absolute) rest" is logically impossible. Since a "resting system of reference" would belong to this class of objects, we suggest accepting as well that such a system is logically impossible.

Because of the validity of points (a) and (b) we suggest accepting the *result* of the mathematical evaluation of the equivalence of figures 1 and 2, *which bears* the consequence of the impossibility of the existence of a resting system of reference (for the reason that other wise the existence of exceptions within logic would have had to be accepted) as objectively correct (true) in respect to the existing universe.[1]

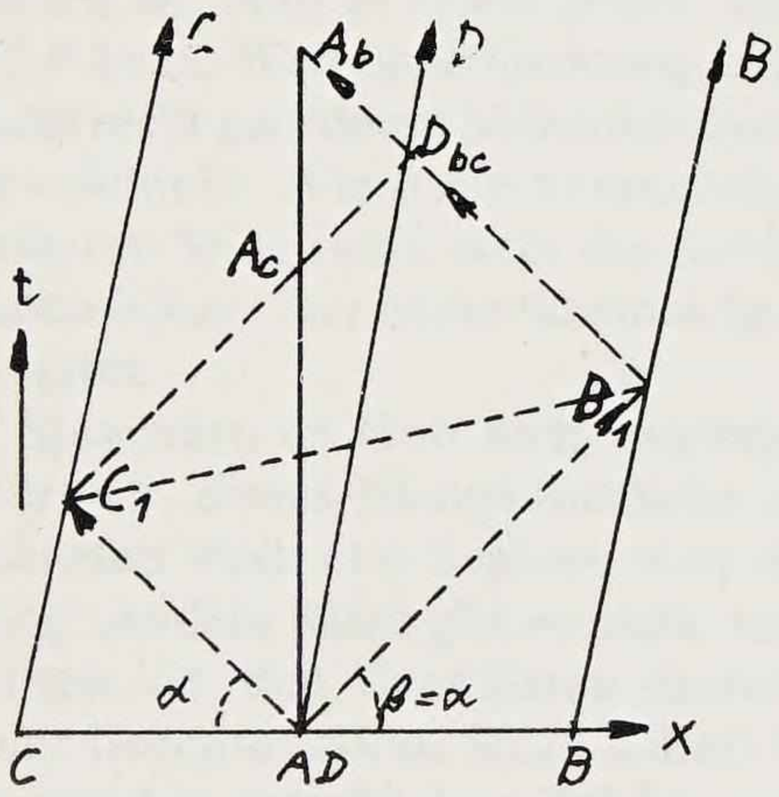


Figure 1

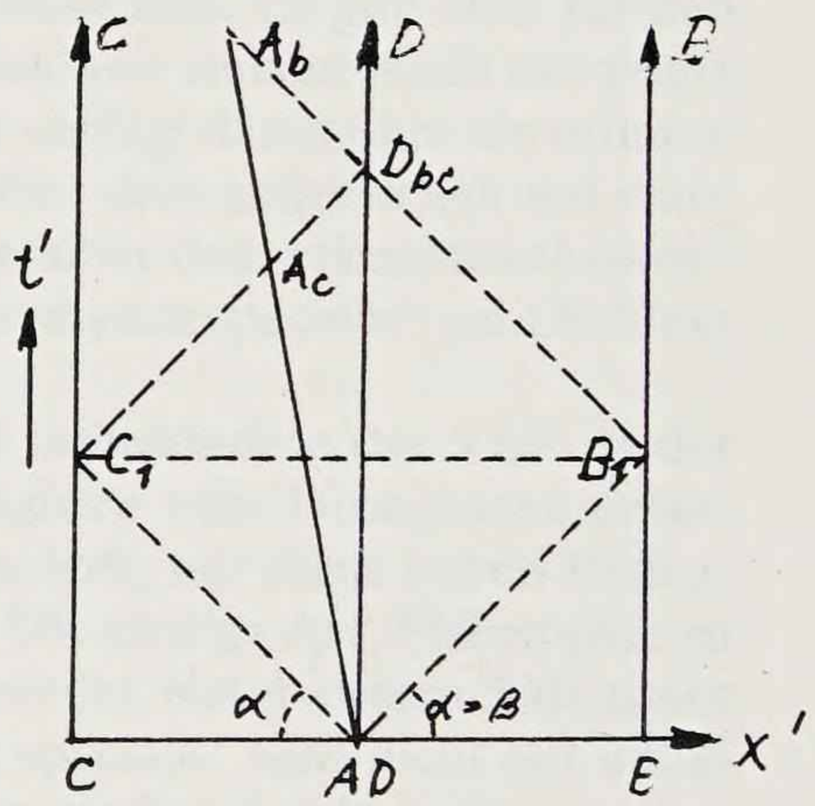


Figure 2

The result of the mathematical evaluation of figures 1 and 2 is $E = mc^2$. We suggest accepting this result as not-theoretically, but as a logical condition of a universe, which possesses the property „to be in motion".

The *verbal* evaluation of figures 1 and 2 (this is the evaluation of figures 1 and 2 by means of an application of "common sense" on the basis of universal regularities of natural languages—which we want to call "conceptual logic") permits certain propositions. If we want to proceed in this process of recognition of our world in a rational way, *we can't help* but apply the intellectual faculties (mind, reason, common sense, "ratio") as common sense by making use of the formulating capacity (the "medium") of the *given* structure of a natural language.[2] (If intellectual faculties working as common sense in the way described above, were not only

incomplete but also "unsuitable" respecting the reason of nature, every rational enterprise having the aim of knowing something, which definitely corresponds to nature's (logical) structure would be absolutely hopeless.–Figures 1 and 2, with their correctness as well as their "world-concern" including how they bear on a correct world-concern had then to be seen as a "strange unlikelihood" as far as "world-suitability" is concerned.)

If we apply common sense to evaluate figures 1 and 2 verbally, we arrive at the following propositions:

1) Space and time are in a state of coupling.

2) There does not exist such a thing as an "absolute time". The times of systems of co-ordinates, which move at whatever different speed against on another are not identical.

3) The transformation of a system of co-ordinates x to a system of co-ordinates x' is necessarily connected with a change of the measure for length, which *can* (and actually does) *take place gradually*.

We suggest accepting as objectively correct both the mathematical as well as the conceptual evaluation of the two figures–where *the "special case" of the existing universe is concerned.*

Since the existing universe is *one* universe (and not universe*s*), it cannot be a) a non-Einsteinian "old" world, obeying the laws of Newton's mechanics, and b) an Einsteinian world, obeying the *laws* of TSR, *at the same time.* If this (one single) universe has to obey *the laws* of TSR, the following law, which stands above all laws of nature, *and which prevents actions of human minds from posing as physical events,* has to be derived from the context of point (3) (*within a process of rational recognition of this world*):

If gradual changes of measures belong to the competence of an intrinsic reason in nature, and are performed as real events within physical reality, abstract gradual changes of measures through the human mind must remain without consequences in a definite way.

If, within an Einsteinian world, which possesses the property "to be in motion", the quantities of all possible types of fundamental physical entities (the fundamental velocity, the elementary mass, the elementary action-quantity, the elementary electric charge) do not depend on measure-conditions of all possible kinds of measures, whereas all these quantities do depend on one another numerically (!), it follows

a) that we will *always* arrive at (numerically) the same quantities for fundamental physical entities independent of measure-conditions, and

b) that existing constant, relations between the quantities of fundamental physical entitites *can be founded on laws only, the character of which is purely formal.*

If the law, which we have deduced from the context of point (3) is a logical condition of the world, its formal representation should, in principle, fall within the range of man's intellectual faculties. If this representation is possible, it should have the form of a geometrical construction, in which the numerical sizes of fundamental physical entities appear as *scalar ratios*, pointing at one another–thus demonstrating that they all are mere *aspects* of a physical reality, which is one formal wholeness. Whithin this construction the measure-unit "one" for length must not play a role at all. It can be seen easily that the length "one" of the radius of the unit circle can be *chosen arbitrarily* and *remains without influence* on goniometric functions. We have good reasons to assume that the formal representation of "our" law should have to do with angles on the basis "r" of the unit circle. Only if we call the radius of a circle "one" and take it as cathetus of a rectangular triangle, the length of the counter-cathetus corresponds to the tangent-value of a given angle. The logic of all goniometric functions contains the *allowance* to choose the length of the radius of the unit circle arbitrarily, but it also contains a prohibition on expressing the size of the radius of the unit circle by a number, which is *not* 1. But here we also have to note that triangles belonging to the sinus- (or cosinus-)form of triangle-constructions are under artificial coercion to remain within the unit circle, while triangles belonging to the tangent-form of triangle-constructions can grow "in a natural way along the numerical series". Depending on the length of the counter-cathetuses, the hypotenuses of such triangles grow in a natural way as well. We are led to the hypothesis: If the natural behaviour of physical reality depends on the natural behaviour of formal elements,

while it cannot be influenced by measure-conditions, goniometric functions of the tangent-form could possibly be the adequate logical framework to demonstrate this dependence.

II

From the recent history of physics we know that two theories remain in a state of rivalry until unified into a single homogeneous and comprehensive description of physical reality. The two "theories" are:
1) The Theory of Special Relativity, within which the speed of light plays an important role.
2) The Theory of Quantum Mechanics, within which the relation $\lambda = h/p$ is significant, but *not* velocities close to the speed of light. If it is presupposed that the result of the TSR ($E = mc^2$) is correct beyond doubt, while quantum mechanics is correct at least to a certain extent, we will suppose that both theories are different but *logically equivalent* descriptions of physical reality. The initial aim of this paper is to show that there is a logical equivalence between the two "conflicting" theories. The primary aim is to show that the title of this paper is significant and true. Our first aim confronts us immediately with two essential questions: a) How is the methodological procedure adequate and how does it promise to attain our goal? and b) which form of proof can we arrive at eventually (in case of "success") to demonstrate that our *understanding* of the logical equivalence of the two theories is correct? Dealing with question (b), first, we have to see clearly that the results of a process of understanding based on facts found by experiments or derived from the evaluation of experiments cannot be subject anew to experimental proof. Thus we are fully aware that our results can—at best—have the form of a geometric-arithmetic construction, showing that the fundamental element of QM, namely *h* (Planck's Constant), "depend on", "leads to", "is a consequence of", "is conditioned by" (etc.) elements of the TSR, as c^2 and (a special amount of) "mass" *within the logic of mathematics*. We have to leave it to the judgement of the scientific and philosophical community, whether "consistently presented connections" between elements of the TSR and QM (in the way as we were able to present them) are accepted as *criteria of objective correctness* or not. The predominant paradigms of both, physics and philosophy will, as we assume, be grave obstacles for an agreement to our presentation.

As far as the methodological procedure (question (a)) is concerned, we want to treat it philosophically in the sense of M. Schlick[3]), making use of the "knowledge-increasing power" of the structure of natural language, this is to say "making use of conceptual logic". We apply conceptual logic to the expression "$E = mc^2$", stating that this expression is not an equation. We justify this statement in the following way: while "mc^2" is constituted by concrete physical facts "E" is an empty symbol, and can as such be replaced by some other symbol. *Only* then, if the expression "E" is *replaced* by a concrete description of physical reality, logically *equivalent but not identical* with the description "mc^2" do we have an equation in front uf us, *which is not tautological.* If possible, we want to show that "QM *completes* TSR" and vice versa. For this process of showing, we cannot see another possibility than equating "somehow" the leading element of QM "*h*" with "mc^2".

For reasons which we have given in previous writings[4, 5]—and which are "outcomes" of another application of pure conceptual logic—the expression "$E = mc^2$" should, as we are convinced, be transformed into "$h^2f^2 = m_ec^2$" to become a non-tautological equation of a synthetic (knowledge-increasing) character. Since we take it for certain that the human mind cannot deviate from the reason of nature and b) that the existing universe is a singularity (and not universes), every kind of *conceptual* dualism with reference to this universe appears to us to be untenable. Presupposing that we are able to arrive at a "final" comprehensive registration of the mathematical content of QM (in the same way as we take "mc^2 as a comprehensive final result of the TSR), the result of an equation of these two final entries should necessarily be a complete presentation of the mathematical foundations of physical reality. To justify the special form of the equation $h^2f^2 = m_ec^2$—and especially to counter the objection that it is a "violation of dimensionality"—we have to lay special emphasis on the following points:

1) Since h is an "action-quantum" with the dimension gcm^2/sec (= impulse x cm *or* energy x sec), it can lead (as it is shown by QM) to permanent states of energy ("electromagnetic vibrations, moving at the speed of light"), which we call "photons" or gamma quants (γ). The energy of photons is *variable* and depends on the frequency ν, according to the equation $E = h\nu$. We claim: because of the "singularity" of physical reality *all* states of energy, no matter how they may appear, must be established by the one and single elementary *action* h. Within one single universe only one elementary action can exist and therefore all physical states necessarily depend on that elementary *action*. Inasmuch as all states of physical reality are states of energy, their energy is a product of the *action* h. Concerning the fundamental material state of energy "electron", the energy of which is *constant*, ν must get a *definite numerical size*, which we will call "f". Therefore the *minimum* fact "hf" must be given with reference to its energy. The energy of the electron is given—according to TSR—with the expression $E = m_e c^2$ (m_e = mass of the electron). Thereout it follows that an equation of the "minimum" form . . . $hf = m_e c^2$ has to exist. With that form "*dimensionality is already exhausted*". All components of the equation have become *scalar quantities*—and we have to accept this exhaustion, and bear all consequences no matter how "strange" they may appear!

2) The energy of the electron is—according to $m_e c^2 - 8{,}187265 \times 10^{-7}\ gcm^2/sec^2$ (erg.) The energy of a photon, corresponding to the energy of an electron is also $8{,}187265 \times 10^{-7}$ erg, *but according to* $h\nu$. Thus it follows that—since $hf = h\nu_e = 8{,}187265 \times 10^{-7}$ *erg*—a "*particle*-reality" $hf = 8{,}187265 \times 10^{-7}$ *erg,* is *impossible,* because such a reality conceptually *is* a "*photonic*-reality", which *corresponds* to the *particle*-reality "electron" (with reference to energy), but certainly *is not* a particle-reality "electron". The conclusion is: though $hf = 8{,}187265 \times 10^{-7}$ *erg* cannot be a complete correct description of the *material* electromagnetic vibration "electron" and yet hf has to occur in the way . . . $hf = m_e c^2$, *purely mathematical derivatives* of hf *have to complete* the equation . . . $hf = m_e c^2$. The reason simply is that there do not exist other action-quantities than h within this universe, which is one, because all its states obey one and the same order—being one and essentially the same in spite of their "different appearence". (The expression "essentially" does not refer to a "metaphysical essence" but only to the fact, that all states of energy have to be results of the one and single action h. In this respect, h can in fact be called "the essence" of physical being.) "Differences of appearence" of states of energy, which are numerically identical concerning to energy (like e^-, e^+ and γ_e) must necessarily depend on the laws of an *internal* (not dimensional, and not dimensionally appearing) natural mathematical structure. In so far as the structure of physical reality cannot be grasped by direct or indirect observations, it is *internal*. Yet, we have an access to this internal structure, because the fundamental action h can appear externally—establishing reality dimensionally, and thus bringing to our knowledge its *importance as a scalar quantity* within the purely mathematical internal structure of the world.

3) Whether sensual appearance is in favour of an internal mathematical structure of the world or not, *conceptual correctness* has to have absolute ("ontological" or "epistemological") *priority*. The opportunity which is offered to us with regard to a correct understanding of the internal logical structure of the world, depends on the fact that we are able a) to get to know the scalar quantity of h by the way of conclusions out of results of certain experiments, and b) that with our knowledge of the scalar quantity of h the hope is justified, that we will be able to grasp the internal mathematical structure of the world *in the form of a bidimensional static geometric construction, which shows the mutual dependence of physical sizes,* may they be components of the TSR or of QM. (The equation . . . $hf = m_e c^2$ allows us not only to disregard dimensions, but it allows us also—according to the purely formal laws of mathematics—to disregard the "decadic step 10^{-7}.) *We* have to justify by way of a bidemensional static geometric construction that our demand "$h^2f^2 = m_e c^2$" is correct. Given the correctness of this equation, the transformation of a (fundamental) photonic electromagnetic vibration into *the polar material complementarity* of a positive and a negative electromagnetic vibration (positron = e^+, electron = e^-), which occurs as a real event under very specific conditions, can be described as follows:[6]

$h\nu = h\,(2hf^2 + x) \rightarrow 2\,h^2f^2\,(e^+ \text{ plus } e^-) + hx \text{ (kinetic energy)}$

4) As we have pointed out previously,[7] "nature" has to be seen as a "natural-logarithmic-mathematician" acting automatically (*without* the possibility of *error*) in the manner of human mathematicians—since there cannot *exist different ways* to do mathematics as ontology. (All human mathematicians, whenever and whereever they may appear in the cosmos need not necessarily arrive at a decadic system of numbers, but necessarily must arrive at an *e* (basis of the system of natural logarithms) logically equivalent to our decadic one.) Just to illustrate the utmost simplicity of a logarithmic view of the logical structure of the world, we will give the following example: ($c_o = 2{,}997925 \times 10^{10}$ cm/sec//10^{10} cm/sec; $1/m_{eo} = 1{,}097748 \times 10^{27}$/g// 10^{28}g;

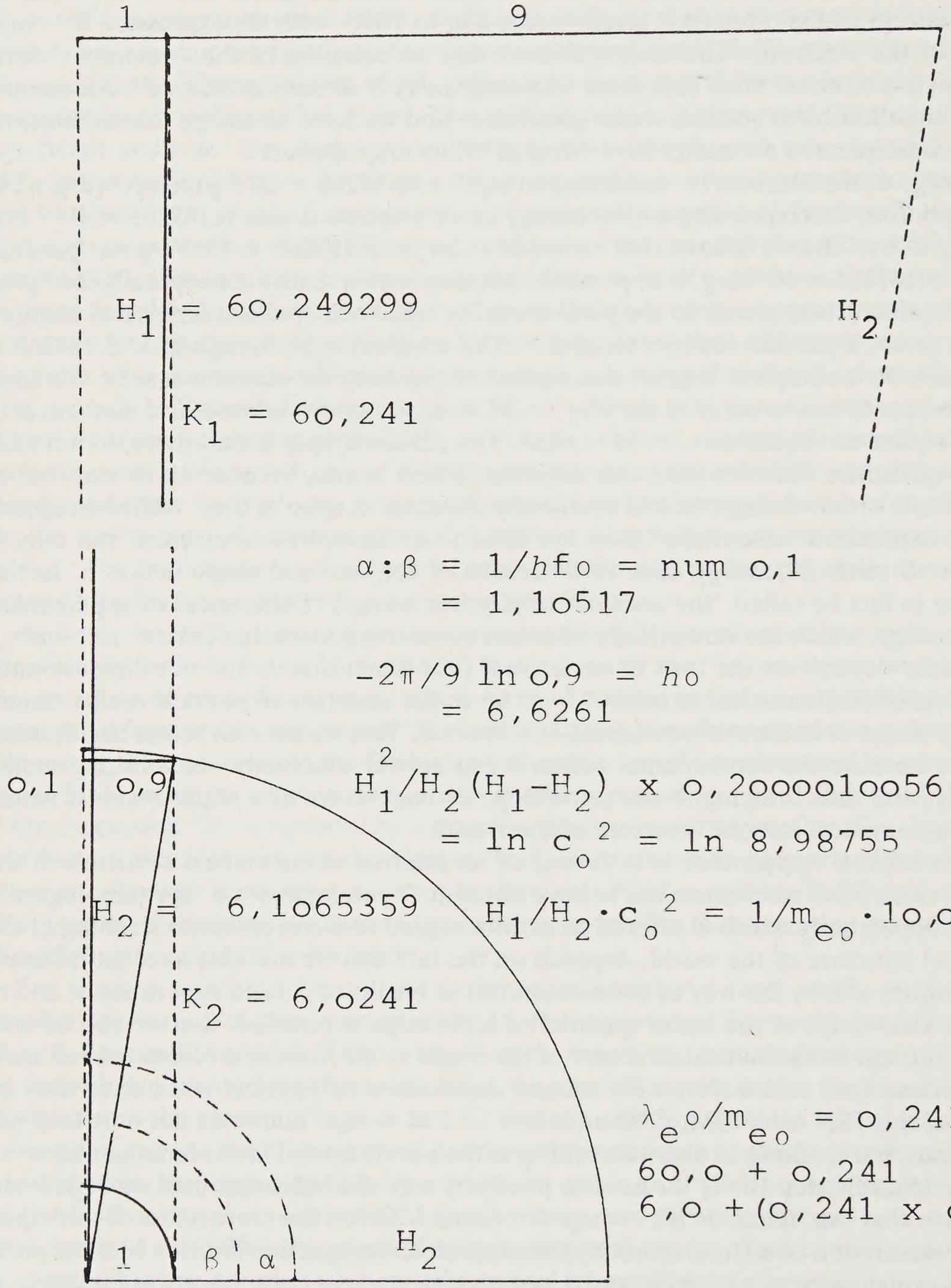


Figure 3

$0{,}1 = \ln (1/hf_o) = \ln (1{,}1051709 \times 10^3//10^3)$:
$\{(\ln c_o) + 0{,}1\} \times 2 = \ln (100 \times 1/m_{eo})$

We see that within a logarithmic view seemingly arbitrary values of components of physical reality remain numerically constant, while only decadic steps and the factor 2 do occur.

These components derive from both, TSR and QM. With this paper we want to to prove that *the human mind and the reason of nature are in perfect agreement* with the mere *gradual* restrictions a) that the human mind has the possibility of making (not easily dedectible) *conceptual* mistakes and b) that the reason of nature does fully understand the complexity of CODES (like *e*'s), which it uses to establish physical reality, while the human mind has *difficulties* in understanding the logic of codes. For the time being, we will confine ourselves to the presentation of three more figures, trying to make clear, in which way *we* understand *the mathematical foundations of physics*.

Figure 3 shows that $1/hf_0$ (as it follows out of the equation $h^2f^2 = m_ec^2$) is the ratio of two distinguished angles $\alpha = 89{,}048979^0$ and $\beta = 80{,}574852^0$, given through the counter-cathetuses 60,241 and 6,0241 at the unit circle (which is the circle with the radius 1). The ln of $1/hf_0$ is 0,1 (numerus $e^{0,1} = 1{,}1051709 = \alpha/\beta$). There is no doubt that for our decadic *e* the exponential step 0,1 is the first (and therefore distinguished) step to establish a relation to the decadic realm of numbers outside of itself. Figure 3 shows that the geometrical representation of the distinguished number 1,1051709 (being the numerus of $e^{0,1}$) depends on a *distinguished "origin"* of the general ratio 10 : 1, and this distinguished origin, given by the ratio of the numbers 60,241 and 6,0241, does exist in the field of mathematical logic *only one time*.

What has to be proved is our assertion that we had to arrive at those number-sequences for fundamental physical sizes, at which we have arrived (independent of the "this-like sizes" of our measure-units) for the reason of an a-priori-dependence of physical reality on the purely formal laws of mathematical logic within the framework of the decadic system of numbers, which we have introduced (see page 182). If we call the hypotenuse of the triangle, given by α, the cathetus 1 and the counter-cathetus 60,241, H_1, and designate the hypotenuse of the triangle, given be β, the cathetus 2 and the counter-cathetus 6,0241, as H_2, we can establish the relation

$$(H_1)^2 / H_2 (H_1 - H_2) = 10{,}979149$$

And if we feed our computer with the numbers of c_o (this is $2{,}997925 \times 10^{10}$ cm/sec$//10^{10}$ cm/sec) and then simply push the button "ln" we get 1,0979204. Hence it is shown that a ratio 1 : 10 exists between the number-sequence of the ln, obtained out of the number-sequence of an empirical measured value and an a priori (formal) value, which follows out of a geometrical construction, in which the distinguished origin of the ratio 10 : 1 is the ratio of two counter-cathetuses at the unit circle. (As one can see, we have expressed the numbers 60,241 and 6,0241 as multiples of the length "*one*" of the radius of the unit circle and have put them into the position of counter-cathetuses.) The connecting lines between the final points of the stretches 60,241 and 6,0241 and the central point of the unit circle produce angles the ratio of which can be understood as the numerus of $e^{1:10}$. (Remark: Just by chance we have learned that the ratio of the specific gravities of "heavy water" and water is 1,105 . . . also.) But the ratio 10 : 1 appears twofold in another way within figure 3: While at its final point H_2 has reached a distance 1 from the vertical line, raised through the centre of the unit circle, H_1 has reached a distance of precisely 0,1 at that height. At the height of the final point of H_1, this hypotenuse has reached the distance 1, while the prolongation of H_2 has reached a distance 9 from the vertical.

If we admit that the way in which physical events occur (the motion of photons occurs at the speed of light) depends on formal laws shown by figure 3, then we come to see that "abstract changes of the measure for length—performed by the human mind—"cannot influence the existing numerical relations: while we can very well choose an arbitrary size for the measure of

length (and can call it centimeter, inch, or foot) we can't help but call the radius of the unit circle 1, and the counter-cathetuses 60,241 and 6,0241.

On the other hand, our construction in two-dimensional space provides us with an a priori knowledge of number-sequences, which correspond (with an accuracy of 5 figures) to empirical values. In addition a mutual dependence of number-sequences, corresponding to empirical sizes of quite different physical entities can be deduced:

$0{,}20000100563 \times (H_1)^2 / H_2 (H_1 - H_2) = \ln c_0^2$ $(c_0^2 = 8{,}98755 \times 10^0)$
$\{(\ln c_o) + 0{,}1\} \times 2 = \ln (100 \times 1/m_{eo})$

If we take the numerically constant difference of the deviations of H_1 and H_2 from the vertical at certain points as meaningful (this difference is 0,9 at the final point of H_2 and 9 at the final point of H_1), we arrive at the number-sequence of Planck's Constant "heuristically": (Heuristics was always a method of mathematicians.)

$1/9 \ln 0{,}9 = -1{,}054802$; $-1{,}0545802 \times -1 \times 2\pi = 6{,}6261227$ $(h_0 = 6{,}626196)$

If we accept that all physical events have to leave the fundamental constants and their relations untouched, we can understand them as rotations of the elements of fig. 3 around the centre of the unit-circle.

While, if we take the ratio $\alpha : \beta$ a thousand times, we arrive at 1/hf, as it follows out of $h^2f^2 = m_ec^2$, we arrive *at the distinguished* character of the 360°-graduation, if we take a thousand times angle α, relate the resulting angle $\gamma = -51{,}022018°$ to $\alpha = 89{,}048978°$ (our computer subtracts automatically 247 complete rotations) and multiply by 2π:

$2\pi (\gamma : \alpha) = 2\pi (-51{,}022018° / 89{,}048978°) = -3{,}6000502$

If the 360°-division is logically distinguished, it follows that codes expressed within this graduation can point directly at numerical physical facts (see figure 4). −3,6 can be understood as "-360_0", like c_0.

(If we use another angle-graduation, e.g. the 400°-graduation, we come to the same result). What remains constant is the ratio of the rotations of α and β.

The number-sequence 241, which we have attached to the lengths 60, and 6,0 is identical with the ratio of the wavelength of the state of energy "electron" (as it is given according to $h^2f^2 = m_ec^2$) to the elementary mass:

$\lambda_{eo}/m_{eo} = 2{,}1954104 : 9{,}109558 = 0{,}2410007$

What is strenghtened by figure 3, is our "conceptual expectation" that the natural mathematical order can—untouched by all arbitrary measure-conditions—only *point back* to itself (showing its different aspects) and cannot deliver *any* nonsensical numbers, belonging to "an unreal transcendent world". Nature as the original mathematician uses "our" system of coordinates, as it uses the system of natural logarithms and the 360°-graduation of angles. We have to accept that space possesses a *discontinuous* geometric-arithmetic structure of a unique form, which presents itself, in the way it does, because of the introduced decadic system of numbers.

Remarks:

1) The "natural size" of $1/hf = 1000\alpha/\beta = 1{,}1051709 \times 10^3$ can be understood from the construction of figure 3 in the following way:
$1/\{\ln(1000\alpha/\beta) - 7\} = \gamma + 180°$ (7,000002017 precisely)

2) The fundamental character of the relation $-2\pi/9 \cdot \ln 0{,}9 = h_o$ can be illustrated, if we write $2\pi \sqrt{1{,}23456789}\ \{10 \ln \sqrt{1{,}23456789}\} = 6{,}6261227 = h_o$. (Note: 1,2−3−4−5−6−7−8−9)
(The value of the "empirical" h, as it is given in recent publications is $6{,}626196 \times 10^{-27} gcm^2/sec$.)

3) The number-sequence 241 appears as a value in diffraction experiments, as they were performed with accelerated electrons by George P. Thomson for the first time. The sharpest diffraction-pictures are obtained, if one uses a tension of 150,52 V. The velocity v of the electrons is according to

$v = \sqrt{2 \cdot e \cdot V/300 \cdot m_e} = 7{,}2739 \times 10^8$ cm/sec ($V = 150{,}52$ V)

Within that formula the expression

$150{,}52 \times 4{,}80325 \times 10^{-10}/300 = \mathit{2{,}41 \times 10^{-10}}$

appears. In our understanding of physics "sharpness" (exactness) depends on exact numerical fit. If the external conditions are made to fit the internal structure, we will get exactness, *which can be "seen"*:

$\lambda_e/m_e \times \mathit{1 \times 10^{-24}} = 150{,}52 \times e/300$ (e = elementary charge)

In Thomson's experiment the numerical correspondence is given twice:

The velocity of the accelerated electrons fits the ratio h/m_e also:

h/m_e : v = 7,2739 × 10° cm²/sec : 7,2739 × 10^8 cm/sec = $\mathit{1 \times 10^{-8}}$cm

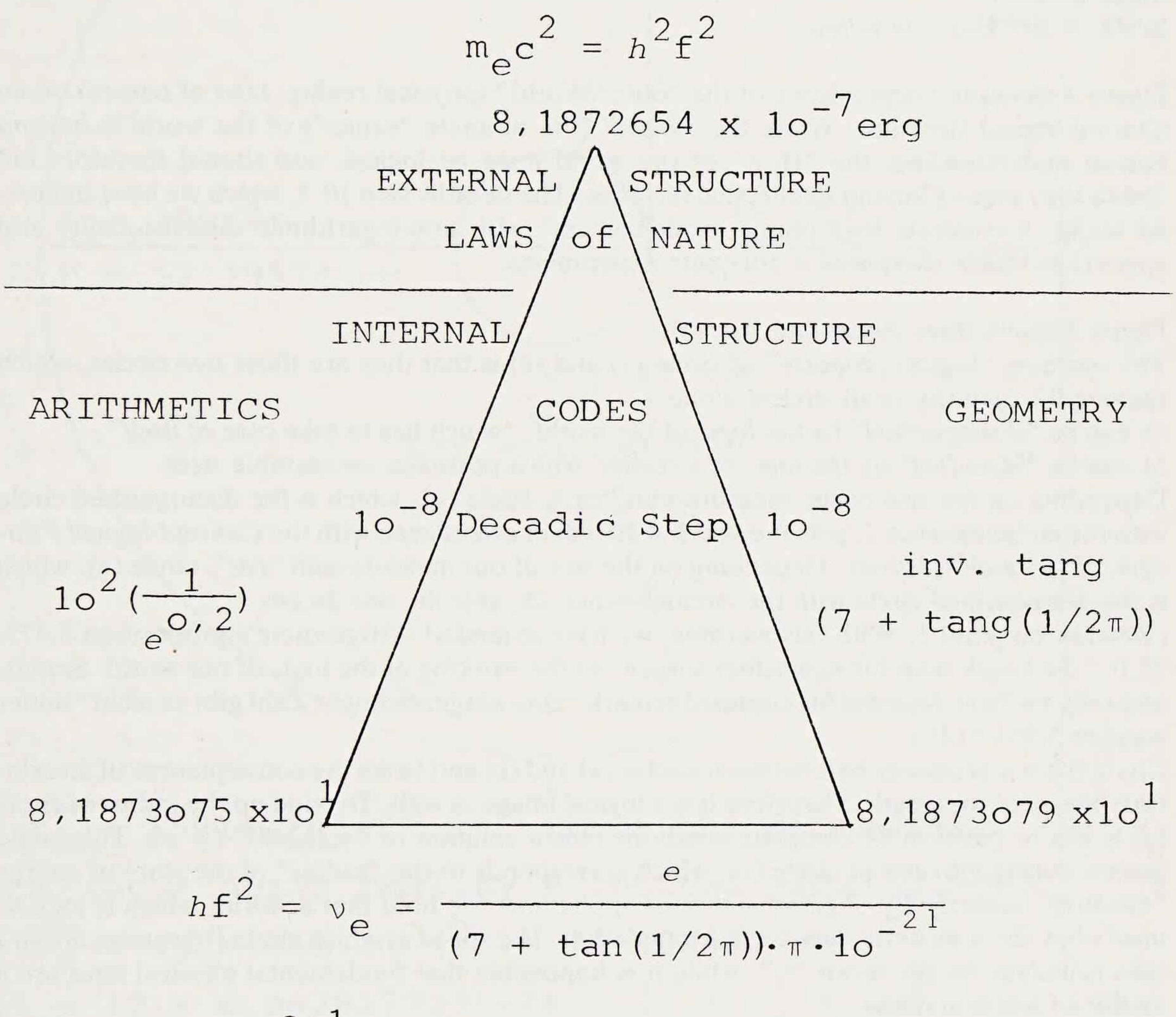


Figure 4

4) To prove and demonstrate the unique and final importance of our "codelike" numerical understanding of nature, our aim must be to deduce fundamental "inertial" values from the equation $h^2f^2 = m_ec^2$.
The mass of the particle "proton" is 1836,1089-times the mass of the electron. We call this value "proton-inertia" ($= 1{,}8361089 \times 10^3 m_e$)
We have found the following relation:
(We will write "$\varkappa$" for $1/hf$ ($= 1{,}1051709 \times 10^3$))
$1 \times 10^6 \cdot \ln(\varkappa/\pi)/\{\text{num}(\ln \varkappa/\ln \pi) \cdot \ln \varkappa\} = 1836{,}1099$
5) In pursuing our goal to show that we do not present a theoretical construction, we succeeded in the following way:
$\ln (c^2/m_e) = \ln (9{,}8660658 \times 10^{47}) = 1{,}105106 \times 10^2$
or
$10{,}000587 \ln (c^2/m_e) = 1/hf$
or (see figure 3)
$\ln (H_1 \times 10^{47}/H_2) = 1{,}1051063 \times 10^2$,
which means
$\ln (H_1 \times 10^{47}/H_2) \simeq \ln (c^2/m_e)$

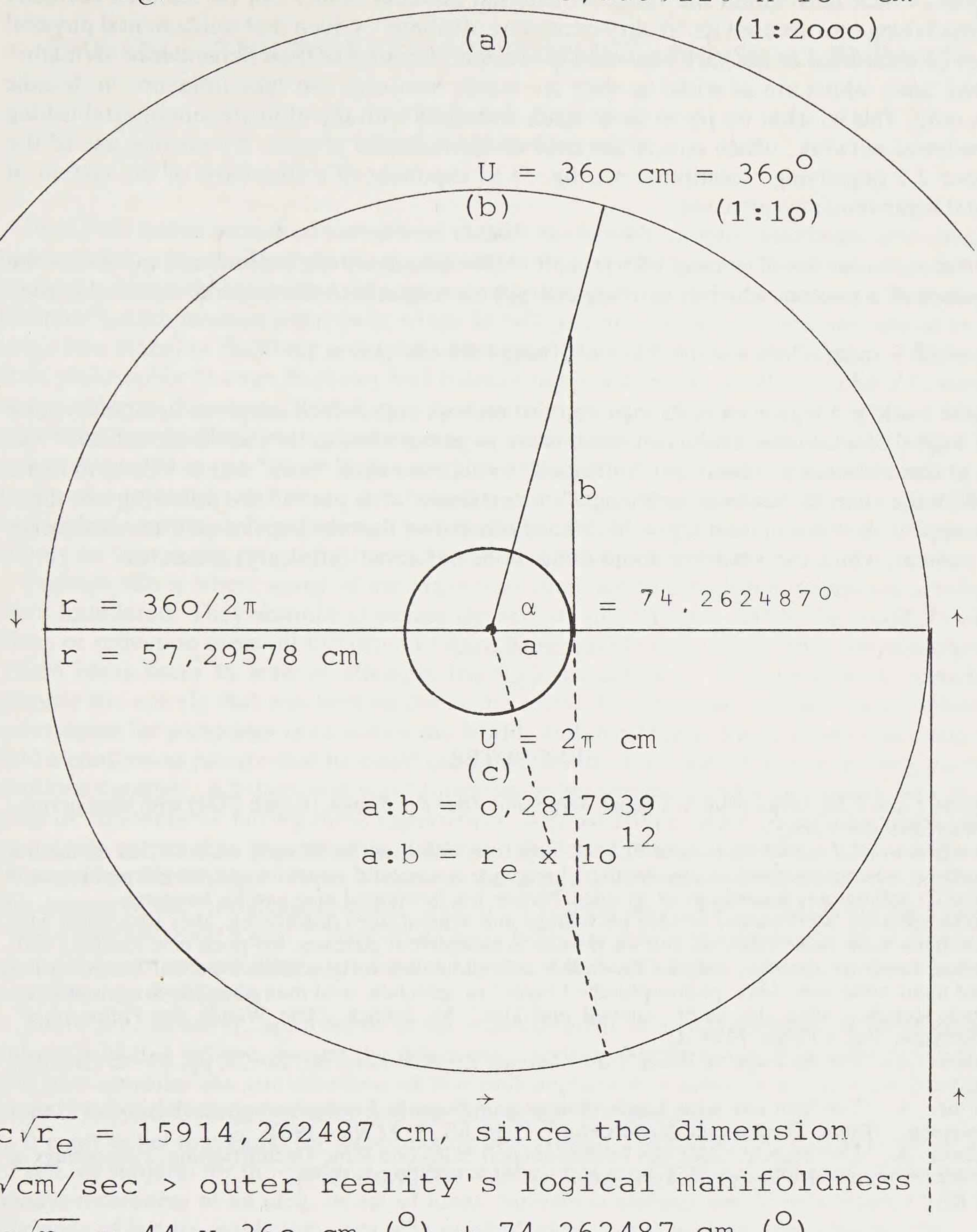
Figure 4 shows the dependence of the "outer World" (physical reality, laws of nature) on an internal logical structure. While the "What" (the ultimate "cause") of the world is *beyond* human understanding, the "How" of the world must be logical, and should therefore fall within the range of human intellectual faculties. The decadic step 10^{-8}, which we have indicated in fig. 4 connects logarithmic "codelikeness" with non-logarithmic dimensionality and appears as visible sharpness in adequate experiments.

Figure 5 shows three circles (a), (b), (c).
The common "logical property" of circle (a) and (c) is that they are those *two* circles, which (among the quantity of all circles) alone
1) can be "*distinguished*" by the *logic* of the world, "which has to take care *of itself*",
2) can be "*identified*" by the logic of a reality, which possesses measurable sizes.
Depending on the size of the measure-unit "cm", circle (a), which is the distinguished circle with the circumference 1, gets the size 1×10^5 cm in accordance with the *Concept-logical Principle of General Relativity*. Depending on the size of our measure-unit "cm", circle (c), which is the distinguished circle with the circumference 2π, gets the size 2π *cm*.
(Remark on point 1: With this wording, we have extended Wittgenstein's proposition 5.473, TLP, "die Logik muß für sich selber sorgen" to the working of the logic of our world. Simultaneously we have rejected his cramped remark "eine ausgezeichnete Zahl gibt es nicht" under number 5.553 TLP.)
Circle (b) is a necessary link between circles (a) and (c) and bears the consequences of an arbitrary measure for length. Therefore it is a logical image as well. To wind up the radius of circle (a) it has to perform 44 complete rotations plus a rotation of $74{,}262487° (= \alpha)$. This angle gives a cotangent-ratio at circle (c), which corresponds to the "radius" of the state of energy "electron" numerically. Against sensuous appearance we hold that a world, which is logical, must obey the laws of its *static* logical foundation. The speed at which circle (b) rotates doesn't take influence on the result "r_e", while it is impossible that fundamental physical sizes are a matter of accidentalness.

If we define a measure for mass, we should not forget that "we do not act onedimensionally but take into consideration 'automatically' the reciprocal value of the square of the speed of light", since with mass we size a property of energy. Moreover we make the measure for the property of matter "to be heavy" depend on our measure for length.

Our point of view is just the opposite from the traditional view which understands that our threedimensional world is embedded in a higher dimensionality. We hold that "the internal logical structure of the world" possesses a logical manifoldness, which is larger than that of

$W = (c\sqrt{r_e})\,2\pi \quad \rightarrow \quad U = 0{,}9999226 \times 10^5$ cm



$c\sqrt{r_e} = 15914{,}262487$ cm, since the dimension $\sqrt{cm}/sec$ > outer reality's logical manifoldness

$c\sqrt{r_e} = 44 \times 360$ cm ($^\circ$) + 74,262487 cm ($^\circ$)

$44 = (7 + 0{,}0028179\overline{39}) \cdot 2\pi = (7 + r_e \times 10^{10}) \cdot 2\pi$ rot.

$2{,}2 = 0{,}7002817939\pi$ (essential part of numeric Code)

Figure 5

physical reality as we know it. We assume that of a given dimensionality (like that of c $\sqrt{r_e}$) the part, which falls within the range of external physical reality can be handled correctly without taking into account its "reality-exceeding attribute". Given that fundamental physical values (dimensional or not) are numerically constant, because of their dependence on mathematical laws, which are as scalar as they are static, measures can take influence in decadic steps only. This is, what we try to show again and again with the ultimate aim of establishing a consistent network, which covers the field of fundamental physics. By making use of the number 2,2 (signifying "rotations", see fig. 5) as exponent of *e* (the basis of the system of natural logarithms) we arrive at

$\sqrt{e^{2,2}}/f\,\pi \times 1 \times 10^{-3} = (7 + \text{tang}(1/2\,\pi)) \times 10^{-21}$ (f = frequency of electron) and can define the frequency of a photon, which (regarding energy) corresponds to the material state "electron" by

$\nu_e = e/\{(7 + \text{tang}(1/2\pi))\,\pi \times 10^{-21}\} = e/\{(\text{tang}\,(100/e^{0,2}))\,\pi \times 10^{-21}\}$

If the world is a logical case, its logic can rest on laws only, which are given by numbers and their logical connections. These connections are as geometrical as they are arithmetical. "The laws of dimensionality" cannot get "infringed" by us, since such "laws" do not belong to logic. If within the context "rectangular triangle" the cathetus "a" is put "1" the following two theorems appear: $h \times p = q$, and $q/p = b^2$. Hence it is shown that the logic of mathematics works as a context, which cares neither about dimensions nor about (arbitrary) measures.

ENDNOTES

[1] Figures 1 and 2 are taken from S. Flügge, *Wege und Ziele der Physik* (Berlin 1974) with kind permission of Springer-Verlag.

[2] We agree to TLP 5,5563 and reject PI § 97. Together with L. v. Bertalanffy we hold that all natural structures must correspond to one another. Language is a natural structure and we learned to use it correctly without any knowledge of its rules. Nature has developed man *and* his language.

[3] "Philosophische Schriftsteller werden noch lange alte Scheinfragen diskutieren, aber schließlich wird man ihnen nicht mehr zuhören, und sie werden Schauspielern gleichen, die noch eine Zeitlang fortspielen, bevor sie merken, daß die Zuschauer sich allmählich fortgeschlichen haben. Dann wird es nicht mehr nötig sein, über 'philosophische Fragen' zu sprechen, weil man über *alle* Fragen philosophisch sprechen wird, das heißt: sinnvoll und klar." M. Schlick, "Die Wende der Philosophie", *Erkenntnis*, Vol. 1 (Wien 1930/31).

[4] Hübner, A., "On the Logic of Being", *International Logic Review*, Nr. 23–24, pp. 27–45 (Bologna 1981).

[5] Hübner, A., "Das Sein und seine Logik–Einige grundlegende Ergebnisse sprachanalytischer Naturforschung", *Wiener Tierärztliche Monatsschrift*, Hefte 6/7, 8, 9 (Wien 1982).

[6] Hübner, A., "Der logische Code des fundamentalen Seins und seine Dechiffrierung, *Proceedings of the 4th International Wittgenstein Symposium*, pp. 585–609 (Wien 1979).

[5] see (5).

* * *

THE SELF, THE WORLD AND THE METAPHYSICAL SUBJECT

Graham P. Conroy
Portland State University

Jean-Paul Sartre remarked somewhere that all modern philosophy must begin with the Cartesian starting point. That point, the Cartesian cogito, the rational terminus of a chain of methodic doubtings has passed through a strange history from its inception as the French philosopher's Archimedean point (with which he believed he could lever a new universe of knowledge into place) to the many criticisms of it within twentieth century philosophy. The Anglo-Irish philosopher George Berkeley had claimed in his notebooks, or *Philosophical Commentaries*: "Cogito ergo sum. Tautology." Descartes' opponent Hobbes had stated that "Cogito ergo sum" was an illegitimate conclusion from general unverifiable premises. Gilbert Ryle had asserted it leads to the "ghost-in-the-machine" syndrome (a spiritual mechanism working the levers of a material mechanism over which it presides). Other assorted linguists, positivists, and logical empiricists have believed with Berkeley: "Cogito ergo sum. Tautology." There are simply no convincing proofs for the existence of self within the world.

Perhaps this is where many of the arguments intended to establish a metaphysical subject have foundered. They attempt to situate the self, or subject, as a part of the world. Locke's ideas of reflection seem all too often to have been used literally rather than metaphorically. These ideas seem to lead to attempts (through inspection or introspection) to somehow glimpse the eye (I) that was looking out at the world. Locke conceived mind as an unknown substratum for picturings seen within the world; and, for Hume, his introspection could not find a continuant picture that he could call mind, a sort of metaphysical glue holding our perceptions together. All there was was "congeries of perceptions" and a metaphoric way of talking of "someone's" having these experiences. Against Hume, Kant placed the world partly in the self, rather than the self in the world and made the mind constitutive of the world in his 'Copernican Revolution'. From Vienna in the *Analysis of Sensations* Ernst Mach celebrated the demise of the "unsalvageable self" and in his Leipzig laboratory Wilhelm Wundt pursued "psychology without a soul." They doubtlessly held that the early empiricists' try at psychologizing the subject and Kant's attempt to logicize it had both failed. There remained just the world and its descripton.

There are, however, approaches to the establishment of a philosophical, or metaphysical subject which, I believe, escape the destructive devices of both rationalists and empiricists. I will now consider the contributions of two philosophers to a solution of the egological problem. They are George Berkeley and the Tractarian Wittgenstein.

It is my belief that Berkeley was breaking new ground. He had substituted a relational idea of self, or subject, for the qualities-substance inherence model of the Lockeans. Awareness is always awareness of an idea, or set of ideas. Subject is unitary, not to be confused with the divisions of faculty psychology, nor with psychological states, nor with dispositions. These are the *anlagen* of self, not the self. They belong to the world side of the relation. I make the claim that for Berkeley the philosophic subject is forever a subject for experience, never an object of experience. Experience happens in the world. For Berkeley "to be, is to be perceived" is the formula for the existence of the world I encounter; "to be, is to will or to act" is the formula for the existence of that which knows the world. Yet strictly speaking we should not call this the epistemological subject. Psychological description is part of the description of the world. Berkeley only speaks of self as will, other than that it is basically uncharacterizable. This view is consistent throughout the *Principles*, *Three Dialogues* and *Alciphron*. Berkeley has often been called mistakenly a solipsist; a subjective idealist, but he is now being seen for

the realist he actually is. The late Professor A. A. Luce often and eloquently made the claim that for Berkeley there was the world of ideas and the self that stood outside them. The body and the sense organs belong to the world of ideas. In a passage which could bear equally well on imputation of both solipsism and realism to Berkeley, Wittgenstein says: "Here it can be seen that solipsism, when its implications are followed out strictly accords with pure realism. The self of solipsism shrinks to a point without extension and there remains the reality co-ordinated with it." (*TLP* 5.64) "Thus there is a sense in which philosophy can talk about the self in a non-psychological way." (*TLP* 5.641) The subject does not belong to the world: rather, it is a limit of the world.

As if to underscore this interpretation Berkeley explicates the phrase 'in the mind' not in terms of a place or a state, but in terms of a relation.

> This perceiving active being is what I call mind, *spirit*, *soul*, or *myself*. By which words I do not denote any one of my ideas, but a thing entirely distinct from them, wherein they exist or, *which is the same thing*, [italics mine], whereby they are perceived—for the existence of an idea consists in being perceived. (*Alciphron*, sect. 2)

Berkeley wanted to say some things about the self but realized the difficulty. He hit upon the term 'notion' as a way out. He said of the notion, not the idea of self:

> Though it must be owned at the same time that we have some notion of soul, spirit and the operations of the mind, such as willing, loving, hating—inasmuch as we know or understand the meaning of these words. (*Alciphron*, sect. 5)

Ludwig Wittgenstein comes off better by saying we can't say anything of the self or subject. It manifests itself. Language is the limits of my world. We may speak only of the world. As a thought experiment he imagines himself to be the author of a book called, *The World I Found*. He states in the *Notebooks of 1914–16*, entry for 23.5.15:

> In the book "The World I Found" I should also have to report on my body and say which members are subject to my will, etc. For this is a *way of isolating the subject*, or rather of showing that in *an important sense* there is no such thing as the subject; for it would be the one thing that could not come into the book. (Cf. *TLP* 5.631)

What Wittgenstein does is to show that the metaphysical subject cannot be part of the world, but must stand in a relation to it; or, if anything can be said about it, that it is the limit of the world. In this act of isolating the self, solipsism, idealism, and realism come into harmony. He writes:

> This is the way I have travelled: Idealism singles man out from the world as unique, solipsism singles me out, and at last I see that I too belong with the rest of the world, and so on the one side *nothing* is left over, and on the other side, as unique, *the world*. In this way idealism leads to realism if it is strictly thought out. (*Notebooks*, entry for 20.10.16)

Now, what can be said about will and its confronting of world. He says among a series of affirmations:

> That life is the world. (Cf. *TLP* 5.621)
> That my will penetrates the world.
> That my will is good or evil. (*Notebooks*, entry for 11.6.16)

The domain of value must lie outside the world. It is for Wittgenstein as it was for Berkeley, and, I believe, Schopenhauer a matter of will. If the world has any value (and for L. W. it does) then that value must lie outside the world. The subject, values, the meaning of life, are all boundary conditions attained at the ladder's top ((Cf. *TLP* 6.54). Beyond that lies the transcendental, the mystical; but mystical only in that it can't be said. Berkeley had called the self a substance, using the substance-talk of the medievals which had been kept as part of the vocabulary of Cartesianism. On Wittgensteinean terms Berkeley is saying what cannot be said. The notion of a boundary conveys the situation more aptly.

It is my belief that the *Tractatus* and the *Investigations* do not contradict each other in regard to philosophy of mind, though the later philosophy takes a great leap forward. By referring to the metaphysical subject as a world limit, and leaving it radically uncharacterizable, it becomes like Musil's man—*ohne Eigenschaften*—whom qualities. This "it" is not a man, however; for as Wittgenstein has stated in the *Tractatus*: "The philosophical self is not the human being, not the human body, or the human soul with which psychology deals, but rather the metaphysical subject, the limit of the world—not a part of it." (*TLP* 5.641) Berkeley would also see body, psychological states, etc., as mind-related, but not as parts of what is meant by person. Subject stands in opposition to its ideas which are the objects of the empirical sciences including psychology. Wittgenstein's explanation is much the clearer despite his aphoristic and elliptical style. He established the metaphysical subject in the *Tractatus* and then explicated his psychological analyses and suggestions in the *Investigations* and the later writings on psychology. In so doing he had already avoided behaviouristic reduction as well as the need for neural identity theories. Life is not just physical, nor just psychological.

> The World and Life are one.
> Physiological life is of course not "Life".
> And neither is psychological life.
> Life is the world. (*Notebooks*, entry for 24.7.16)

Approaching the question of the self from different directives, both Berkeley and Wittgenstein have transcended the Cartesian starting point and placed themselves outside that tradition and its problems. This must be judged one of the most important achievements of each man.

* * *

THE SELF AND INTENTIONALITY

B. S. Gower
University of Durham

In this paper I discuss a problem revived in recent work by Bernard Williams and Thomas Nagel,[1] namely the problem of successfully relating subjective self to objective reality.

First of all there is the idea that among all the things, events and states of affairs that the world may be said to contain, there is one item which has a special significance. This is what I call the *self*. Events and states of affairs which, in some way, involve this self are real and important. Indeed, from this subjective point of view, it can easily come to seem that nothing but whatever involves the self is either real or important. As Russell put it, in a characteristically vivid manner, there is, for each of us, "a region which is specially warm and intimate and bright, surrounded in all directions by gradually growing darkness."

This view, because it is essentially subject-involving, quite readily lends itself to forms of idealism, scepticism or solipsism. Indeed, these philosophies may be said to derive their considerable strengths from the plausible claim that neither the self nor its experiences provide sufficient resources for illuminating Russell's darkness. So, idealism greets the suggestion that there is any such thing as experience-independent matter, which happens to lie outside the scope of the illumination provided by the self, as dispensable superstition. And scepticism trades on an indisputable consequence of idealism, namely that what may not exist cannot be an object of knowledge.

So if we take, as our philosophical starting point, the self together with its expriences, beliefs and feelings, it is hard if not impossible to find a place for a self-independent reality. The point is most familiar in an epistemological context, but the emphasis I want to press is metaphysical. In taking the subjective view we are in effect, declaring that only what is experienced "from the inside" counts as genuinely real. There is no room in such a view for "objectivity"-thoughts, i.e. thoughts about what cannot be experienced "from the inside", because there is no way of conferring meaning on such thoughts. What is being questioned here is not so much the possibility of knowledge concerning self-independent reality, as the possibility of coherent thought about it. This centred, subject-involving way of thinking leaves no space for realism about the world. Realism about the self is purchased at the cost of idealism about the world.

Because of its connection with standard metaphysical, semantic and epistemological issues, the subjective view has always been prominent within philosophical thinking. No-one with knowledge of post-Cartesian epistemology can fail to notice that it implies that thought about the world must be centred upon human experience because the terms in which we think have meaning only in so far as they are correlated with experience. There is no possibility of intelligible thought, still less of justified belief, about what lies beyond, behind or above experience. Man is the measure of all things, and if, as may seem inevitable, the scope of experience is restricted to that of an individual self, then each person is, severally and individually, the measure of all things.

Secondly, and opposed to this, there is an equally familiar conception according to which there is nothing in the least special about human beings and therefore nothing special about any self. Instead of thinking about reality as it appears from some particular viewpoint within reality, in which case the self which has that viewpoint will have an ontologically privileged role, we think about reality as it is, independently of any viewpoint. Clearly the availability of a conception of reality as it is – an objective conception of reality – derives from our capacity

for detachment, i.e. our ability to step back from the centred, subject-involving way of representing reality.

In exercising this capacity for detachment in order to think about reality as it is, each of us thinks of reality as *containing* the individuals that we are. But in envisaging the world in this way we seem somehow to lose something important. The self, when presented to thought *as* a self rather than as an occupant among others of objective reality, cannot be incorporated into an objective or absolute conception precisely because that conception is not subject-involving. There is, as it were, no room in this conception for "I"-thoughts. How can the self that I refer to when I use the word "I" and that I essentially think about and know "from the inside", be identified with any occupant of objective reality, which is always and essentially something that I think about and know "from the outside"?

The question is not new. In one idiom or another it has been, and is, subject to close and careful philosophical scrutiny. Kant's observation that the "I think" which must accompany all representations is, despite Descartes, no more than a framework for thought, is one way of expressing, and perhaps dissolving, the difficulty. Another is to be found in Wittgenstein's remarks on the self and solipsism in the *Tractatus*. More recently, philosophers have argued that "I" fails to refer to anything at all, that "I"-thoughts resist incorporation into a Fregean semantics, that the self is without content or essence, that there are no first-person propositions. Whatever the particular merits of these conclusions may be, the arguments for them have made it plain that if there is a difficulty about identifying the self with any occupant of objective reality, the fault lies not with our conception of objective reality but with our conception of the self. We assume, wrongly, that the self is some kind of thing which *should* occupy a place in any comprehensive conception of reality.

Our realist conception of an objective spatio-temporal framework, then, entails an idealist conception of the self, or, which is the same thing, that "I" does not refer to anything.

We have, then, two mutually hostile pictures or conceptions. In one—the subjective conception—the central idea is that thoughts, beliefs, perceptions, etc., whatever they may be about, are essentially subject-involving and therefore perspectival. Because of this, realism about the self cannot be abandoned even if it entails anti-realism about everything other than the self. In the other—the objective or absolute conception—the central idea is that thought, belief, perception, etc., must be object-involving. That is to say, thought must, somehow, reach beyond itself to grasp what is independent of thought. Realism about a mind-independent world, then, is not a dispensable option but an essential presupposition of the possibility of thought. It, therefore, must be retained despite its hostility to realism about the self.

There is no question of repudiating one or other of these pictures in an attempt to resolve the tension between them. What is needed, rather, is a way of understanding the issues generating this tension so that its philosophical perplexity may be eased. I shall appeal to the concept of intentionality in order to effect such an understanding.

The objective or absolute conception of reality is expressed in thinking which is non-perspectival, non-indexical. The materials for explaining why such thinking is essential are already available. Thought has an intentional object; it reaches out beyond itself. This is why thought is *object*-involving, and the appropriate way to represent the object that is involved *as* an object is to employ a non-indexical mode of presentation. To eschew non-indexical thought would be to ignore what the arrow of intentional thought points to.

This conclusion is uncontentious. Nobody seriously proposes that non-indexical thought and the absolute conception could be abandoned even in principle. There are though, those who would urge that indexical thought does not require the existence of indexical reality—require, that is, the existence of indexical entities like the self. But since thought is intentional this form of idealism can be resisted. If, in thought about reality, our concern is with what directs the thought—with what makes thought about reality *thought*—then we are bound to employ a subject-involving indexical mode of presentation in representing reality. Arrows of intentionality need tails indicating their origins as well as heads pointing at their targets if the idea of their having direction is to make any sense.

My claim, then, is that two quite general philosophical problems raised by reflection on the self can best be understood in terms of the concept of intentionality. The first problem is the ancient one of showing how, if we begin our philosophy with the subjectively conceived self, room and role can be found for an objective world. The second problem is, in a way, the mirror image of the first. If we begin our philosophy with an absolute conception of reality, no place can be found in it for a real self. How can a picture which is essentially subject-involving become object-involving; how can a picture which is essentially object-involving become subject-involving? The double character of the intentionality of thought helps us to understand these two problems and the tension between the two pictures they present. It can be expressed by saying that any thought about any object must be a thought *of that object* (i.e. object-involving) and, at the same time, a *thought* of that object (i.e. subject-involving). In this way, intentionality helps us to understand the tension between objective and subjective views, but it does not dissolve that tension. Indeed it cannot do so because it shows that tension to be necessary.

ENDNOTE

[1] Williams, B., *Descartes: the Project of Pure Enquiry* (Harmondsworth 1978); Nagel, T., "The Objective Self" in Ginet, C., and Shoemaker, S., (eds.), *Knowledge and Mind* (New York – Oxford 1983), p. 211ff.

* * *

4. Methodologie der Philosophie des Geistes

4. The Methodology of the Philosophy of Mind

METHODOLOGISCHE BEMERKUNGEN ZUR PHILOSOPHIE DES GEISTES

Henri Lauener
Universität Bern

Meine Auffassung von *offener Philosophie* entspringt einem Verständnis der transzendentalen Methode, das nicht zu einer im Sinne eines Empiristen anstößigen Metaphysik führt. Ähnlich wie Kant setze ich bei *Fakten* an, nämlich bei den mannigfaltigen Sprachsystemen und Theorien, die wir – als Erzeugnisse nicht der reinen Vernunft, sondern unserer Praxis – selbst gemacht (*facta*!) haben und deren verschiedene Bedingungen es zu untersuchen gilt. Im Gegensatz zu ihm glaube ich nicht, daß es nur *eine* endgültig wahre, d.h. *objektiv* ausgezeichnete Strukturierung der Welt gibt. Trotz meines Verzichts auf eine Letztbegründung bleibe ich Kants allgemeiner Konzeption insoweit treu, als ich daran festhalte, daß wir die in unserer Erkenntnis wirksamen Ordnungsprinzipien selbst schaffen. Mit dieser an den transzendentalen Idealismus gemahnenden Einsicht verbinde ich allerdings einen weitgehenden Relativismus, der mit Absolutheitsansprüchen, wie sie Kant im Anschluß an den traditionellen Rationalismus erhebt, unvereinbar bleibt. Die Menschen gesamthaft betrachtet oder wenigstens größere Kulturgemeinschaften teilen höchstens das, was man in Anlehnung an Wittgenstein ihre Lebenswelt in der weitesten Bedeutung des Wortes nennen könnte. Von einem solchen vagen, noch kaum systematisch zu erfassenden Hintergrund heben sich die durch vielfältige Zielsetzungen motivierten Unternehmen ab, die jeweils einen besonderen Kontext bilden und sich im Extremfall zu einer spezifisch wissenschaftlichen Theorie verselbständigen. Die Tatsache, daß sich bei einem derartigen Verzweigungsprozeß Grenzen ständig verschieben, scheint mir jedoch kein genügender Grund zu sein, um in holistischer Manier Sprache, Logik, Mathematik, empirische Theorie etc. zu einer einzigen amorphen Masse zu vereinigen, und zu behaupten, daß diese nur gesamthaft vor das Tribunal der Erfahrung gestellt werden könne. Die Rechtfertigung eines solchen Vorhabens würde eine *objektive* Auszeichnung der Quantifikationstheorie erfordern, die aber von einem naturalistischen Standpunkt nicht zu erbringen ist. Ich glaube, daß sich in der Tat keine erkenntnistheoretische Position – selbst nicht die, für einen Empiristen naheliegende, des physikalischen Realismus – sachlich als zwingend oder gar als wahr ausweisen läßt: die Gründe, die uns bewegen, eine Theorie einer anderen vorzuziehen, sind rein methodischer Natur.[1] Deshalb empfiehlt es sich, jeglicher Form von dogmatischem Reduktionismus eher mit Mißtrauen zu begegnen und sich für eine liberale, d.h. pluralistische Haltung zu entscheiden. Welchem Vorgehen jeweils der Vorzug gebührt, wird sich durch erfolgreiche Anwendung des theoretischen Rüstzeugs erweisen, das wir für die Lösung der anstehenden Probleme bereitgestellt haben.

Kant hatte nachdrücklich geleugnet, daß wir die Dinge so kennen, wie sie an sich sind. Ähnlich wie er halte ich dafür, daß wir nicht in der Lage sind, unsere theoretischen Konstrukte mit einer neutralen, unmittelbar gegebenen Realität zu vergleichen. Wohl prüfen wir in gewisser Hinsicht wissenschaftliche Theorien aufgrund von sogenannten Tatsachen; da diese als sprachlich artikulierte Gebilde begrifflich durchsetzt sind, erweist sich das, was wir im Kontext für „Erfahrungstatsachen" ausgeben, immer schon als durch bestimmte Interessen geleitete Interpretation. Der Vergleich vollzieht sich also gewissermaßen auf dem Hintergrund zweier Theorien – einer elaborierteren der Wissenschaft und einer primitiveren der alltäglichen Wahrnehmung oder der Experimentaltechniken. Daraus folgt, daß Wahrheit in einem strengen Sinne nicht, wie naive Realisten behaupten, in irgendeiner Übereinstimmung mit der Wirklichkeit bestehen kann. Indem ich dazu neige, das Prädikat ‚wahr' nur im Zusammenhang mit Fragen der formalen Sematik zu verwenden, verweise ich das Problem der Bewertung von empirischen Theorien in den Bereich der Pragmatik, wo es sich im wesentlichen

darum handelt, ihre Brauchbarkeit relativ zur ursprünglichen Zielsetzung – unter anderem dem Erklären von Vorgängen oder der Voraussage von Ereignissen – zu beurteilen. Es scheint mir im übrigen wichtig, den *normativen* Charakter der zu einem solchen Zweck jeweils herangezogenen Kriterien hervorzuheben, über den die in einem deskriptiven Ton gehaltenen Wahrheitstheorien gewöhnlich hinwegtäuschen. Das bedeutet aber nicht, daß ich der nackten Willkür, d.h. der Methode „anything goes", die Schleusen öffnen möchte. Verzicht auf absolute Gewißheit kommt ja insofern nicht einer totalen Skepsis oder der Anarchie gleich, als auch die jeweils befolgten methodischen Regeln durch den praktischen Bewährungsprozess miterfaßt werden. Die durch Kritik angeregten Korrekturen können zwar auch nicht mehr als einen relativen Erfolg verbürgen, was jedoch für unser Fortkommen durchaus genügt.

Quines extremer These vom rein empirischen Ursprung allen Wissens halte ich den Begriff eines relativen Apriori entgegen, der es mir erlaubt, im Kontext spezifischer Zweige der Erkenntnis eine strikte Trennung zwischen Sprache und Theorie vorzunehmen. Ich habe in früheren Arbeiten[2] ausführlich erörtert, wie wir beim Aufbau von linguistischen Rahmenwerken aufgrund von stipulativen Verfahren Sätze als analytische auszeichnen können, um das zu erzeugen, was man mit einem Unterton des Mißtrauens als „truth by *fiat*" bezeichnet hat. Im wesentlichen handelt es sich darum, eine Konvention einzugehen und dies zwar im Sinne einer Vereinbarung, daß die Wahrheit der betreffenden Sätze nicht angetastet wird, solange der Beschluß, das Sprachsystem aufrechtzuerhalten, andauert. Erst wenn es sich aus einem praktischen Grund als unbrauchbar erweist – sei es, weil es zu ausdrucksarm ist, einen zu geringen Präzisionsgrad zuläßt oder sonst welche Mängel aufweist, – entschließen wir uns, es als Ganzes aufzugeben, um ein neues, den vorliegenden Bedürfnissen besser angepaßtes Instrument herzustellen. Ein derartiger auf einen Kontext und eine Sprache relativierter Begriff der Analytizität scheint mir für die praktisch erforderliche Stabilisierung unserer Darstellungssysteme unerläßlich. Denn ohne solche normierende Eingriffe würde eine angemessene Organisation der Sprache gar nicht gelingen. Analytizität betrifft demnach keine Sachfragen über tatsächlichen Gebrauch oder „begriffliche Notwendigkeit", sondern die Frage nach einer adäquaten Reglementierung: Sätze werden durch pragmatisch motivierte Postulate zu analytischen *gemacht*. Damit verliert das Problem auch den Anschein geheimnisvoller Tiefgründigkeit; als transzendentale Bedingung für die Konstruktion linguistischer Systeme stellt Analytizität nicht länger ein unfaßbares Irrlicht dar, das den Philosophen narrt, sondern das schlichte Ergebnis einer empirisch durchaus zugänglichen Praxis,[3] die keines Rekurses auf Bedeutungen als platonische Entitäten bedarf.

Meiner Methodologie gemäß sind somit die apriorischen Formen, die der Strukturierung unserer Theorien dienen, nicht wie bei Kant durch die reine Vernunft endgültig festgelegt; Konventionen können, nachdem sie sich als unzweckmäßig erwiesen haben, widerrufen werden, so daß trotz vorläufiger Immunitäten grundsätzlich *jeder* Satz revidierbar bleibt. Wenn wir durch den Druck der Umstände gezwungen werden, eine Sprache aufzugeben, nehmen wir durch die Wahl eines neuen Systems mit entsprechenden semantischen Regeln das vor, was ich eine *externe* Revision nenne, während etwa die Verwerfung einer wissenschaftlichen Hypothese und ihre Ersetzung durch eine andere innerhalb desselben linguistischen Rahmenwerkes eine interne Revision ausmachen würde. Ein Wechsel der ersten Art kommt im übrigen der Schaffung eines neuen Kontextes gleich,[4] was eine weitgehende Inkommensurabilität der früheren mit der späteren Theorie zur Folge hat.

Die pluralistische Auffassung, die ich Quines Idee einer umgreifenden und vorläufig besten physikalischen Theorie entgegenstelle, wirkt sich ferner darin aus, daß ich wesentlich mehr Arten von Dingen in meiner Ontologie dulde als der Naturalist. Wir verwenden das Wort ‚existieren' in z.T. radikal voneinander abweichenden Kontexten. Im Rahmen der Arithmetik, wie sie etwa von Peano axiomatisiert wurde, gilt der Satz ‚Es gibt eine Primzahl zwischen 5 und 10' als analytisch wahr, weil er mit Hilfe der Ableitungsregeln aus Postulaten herleitbar ist. Aus meiner Sicht bedeutet das nun, daß wir durch die Verwendung des erwähnten Axiomensystems selbst abstrakte Objekte, nämlich natürliche Zahlen, erzeugen und daß es deshalb vergeblich wäre, solche in der empirischen Wirklichkeit suchen zu wollen. Andererseits

finden wir in der griechischen Mythologie Götter wie Zeus oder in Sir Arthur Conan Doyles Roman *Der Hund von Baskerville* einen Detektiv namens Sherlock Holmes, was zur Folge hat, daß in diesen Kontexten respektive die beiden Sätze ‚$(\exists x)$ (x = Zeus. $\wedge$ x ist ein Gott)' und ‚$(\exists x)$ (x = Sherlock Holmes $\wedge$ x ist ein Detektiv)' wahr sind. Das wird aber nur dadurch möglich, daß wir aufgrund von Tätigkeiten unserer Einbildungskraft wie diejenigen des Erfindens von Mythen, Legenden etc. oder des Verfassens von literarischen Werken selbst fiktive Entitäten ins Leben rufen, die sich einer wenn auch nicht realen, so doch in der Welt der Literatur öffentlich nachweisbaren Existenz erfreuen. Anders verhält es sich, wenn wir von der Sprache einen *strikt deskriptiven* Gebrauch machen: in diesem Falle verfolgen wir die erklärte Absicht, über wirkliche Dinge zu berichten, deren Existenz sich nur vermittelst empirischer Beobachtung ausmachen läßt. Handelt es sich dabei um nicht unmittelbar zugängliche Objekte – etwa um den Nachweis von Elementarteilchen in der Quantenphysik, so sind für das Fixieren ihrer Spuren komplizierte Apparate und Techniken erforderlich. Falls sich aber der Versuch, z.B. mentale Entitäten in experimentell kontrollierbarer Weise einzuführen, als hoffungslos erweisen sollte, so müßten wir uns damit abfinden, daß sich auf solche beziehende Theorien einer effektiven Erklärungskraft entbehren und daher aus dem Bereich der Naturwissenschaften auszuscheiden hätten. (Wie wir später sehen werden, würde das allerdings nicht bedeuten, daß sie *überhaupt* nicht existieren: im gewöhnlichen Alltag, wo es weniger auf streng wissenschaftliche Überprüfung ankommt, hat sich ihre Einführung für den Zweck einer leichten Verständigung durchaus bewährt.)

Meine Auffassung hat nicht eine Trivialisierung der Ontologie zur Folge. Ich behaupte keineswegs, daß alles auf undifferenzierte Art existiert. Ich wähle in diesem Zusammenhang mit Quine die klassische Quantifikation als linguistisches Instrument,[5] weil sie sich für die Reglementierung des Gebrauchs von ‚existieren' insofern besonders eignet, als sie es erlaubt, ein präzises Kriterium für ontische Annahmen zu formulieren[6]. Mein nuancierter und – wie ich glaube – unserer Intuition angemessener Entwurf eines Systems der Ontologie entspricht im übrigen dem transzendentalen Anliegen, indem er den beim Aufstellen von Theorien wirksamen Verfahren der Strukturierung und Individuierung, die sich relativierend auswirken, gebührend Rechnung trägt, ohne in eine dem Empiristen nicht zumutbare Form von Idealismus zurückzufallen. Die Tätigkeiten des Beschreibens und des Identifizierens von Objekten hängen von der Verwendung einer Sprache ab. Es kann sogar geschehen, daß wir mit der Wahl einer solchen gewisse abstrakte oder fiktive Entitäten selbst schaffen, was jedoch genau dann nicht zutrifft, wenn wir ein in meinem engen Sinne deskriptives Rahmenwerk einsetzen. Einerseits kann es in einem Kontext, in dem das Prädikat ‚Elektron' nicht vorkommt, kein Elektronen geben. Andererseits ist aber die Frage, ob die Terme eine Referenz in der wirklichen Welt haben, nicht rein sprachlich oder durch Konventionen auszumachen, sondern muß dadurch empirisch entschieden werden, daß man ein Ding der betreffenden Art vorweist. In meiner Transzendentalphilosophie ist also nicht mehr von einem Vergleich der Erkenntnisobjekte mit Dingen an sich oder von der Abbildung von Sachverhalten in einer mentalen Beobachtungssprache die Rede; es verhält sich vielmehr so, daß Sprache und Realität sich wechselseitig bedingen, indem uns die erstere die für den Akt des Referierens erforderlichen Ausdrucksmittel liefert, während die Frage, ob wir mit diesen etwas Wirkliches bezeichnen, von der Beschaffenheit der letzteren (einschließlich unserer eigenen natürlichen Organisation) abhängt.

Meine methodologisch motivierte Version des Pluralismus betrifft ebensosehr die Zielsetzungen wie die theoretischen Ausformungen der Erkenntnis. Indem sie jegliche objektive Auszeichnung mit Hilfe eines Korrespondenzbegriffs der Wahrheit vereitelt, rückt sie den *normativen* Aspekt der Kriterien ans Licht, die bei der Beurteilung rivalisierender Theorien herangezogen werden. Wenn wir eine Theorie für „wahr" halten, so tun wir das *de facto* mit Rücksicht auf ihre „Güte", d.h. die Möglichkeit, sie erfolgreich anzuwenden. Entsprechend müssen wir uns bezüglich der Ontologie mit der Einsicht in die Relativität von Reduktions- oder Identifikationsverfahren abfinden: da jedes von ihnen bestimmte Regeln („legislation" im Sinne von Quine) und Setzungen („posits") voraussetzt, wird es uns nicht gelingen, ein

System allen übrigen gegenüber objektiv auszuzeichnen. Damit will ich nicht leugnen, daß wissenschaftliche Hypothesen experimentell geprüft werden, sondern nur auf die Tatsache aufmerksam machen, daß das, was wir in einem bestimmten Kontext als Daten der „Erfahrung" hinnehmen, immer schon das Ergebnis einer (primitiveren) Theorienbildung ist.

Die geschilderten Züge sprechen zugunsten einer liberalen Methodologie, die eine flexible Einstellung hinsichtlich der Wahl von Sprachsystemen und logischen Kalkülen empfiehlt. Das Aufstellen einer für empirische Theorien adäquaten Semantik erfordert eine offene Haltung, die der Kontextabhängigkeit der Bedeutung theoretischer Terme, aber auch dem intentionalen Charakter der Referenz und ganz allgemein der institutionalen Seite der Sprache mit ihren konventionellen Typen und Regeln gerecht wird. Auf der vorwissenschaftlichen Stufe der Alltagsrede wäre im besonderen ein Verfahren auszuarbeiten, das uns in die Lage versetzen würde, zwischen Regelmäßigkeiten, die auf das bloße Erlernen einer Sprache zurückgehen, und solchen, die durch das bewußte Adoptieren einer Theorie bedingt sind, zu unterscheiden.[7]

Wie wirkt sich nun des näheren die skizzierte Methodologie in der sog. Philosophie des Geistes aus? Auf intuitiver Ebene haben sich zwei allgemeine Überzeugungen verankert, die unter dem Titel „ontologischer Dualismus" und „mentale Verursachung" ihren Niederschlag in der Philosophie gefunden haben; der ersteren gemäß sind geistige Phänomene ihrer Natur nach wesentlich von physischen zu unterscheiden, während nach der letzteren sich mentale Zustände in Form von Handlungen etwa auf das empirische Geschehen auswirken können. Die logische Spannung, die eine gleichzeitige Aufrechterhaltung beider Thesen nach sicht zieht, ließe sich historisch am Lösungsversuch des psycho-physischen Parallelismus der Okkasionalisten, an der von Leibniz aufgestellten Doktrin der prästabilierten Harmonie oder an Spinozas Einsubstanzenlehre illustrieren. Was neuerdings die an einem physikalistischen Weltbild orientierten Empiristen unseres Jahrhunderts angeht, so neigen die meisten von ihnen dazu, unter dem Druck der Annahme einer kausalen Geschlossenheit des physischen Bereichs, den Dualismus dem Ideal einer Einheitswissenschaft zu opfern. In der Absicht, psychische Ereignisse auf physische zurückzuführen, wenden sie des öftern die von Quine semantischer Aufstieg („semantic ascent") genannte Strategie an, indem sie ihre Aufmerksamkeit weniger auf die psychischen Objekte selbst als auf die diversen Mängel unserer mentalistischen Redeweise richten. Meinerseits erblicke ich, abgesehen vom Fehlen einer eigentlichen Theorie, die Schwäche einer bloßen Begriffsanalyse im Stil von Ryles *Concept of Mind* darin, daß ihre Anhänger – sich mutwillig über die experimentelle Forschung hinwegsetzend – eine Kluft zwischen der Philosophie des Geistes und der empirischen Psychologie aufgerissen haben, die meines Erachtens keine fruchtbaren Ergebnisse gezeitigt hat.

In Opposition zu der aus Oxford und Cambridge stammenden sprachanalytischen Tradition beharrt der – in seiner extremsten Form von Quine vertretene – Naturalismus darauf, daß allein naturwissenschaftliche Methoden zur Konzipierung von Theorien mit echter Erklärungskraft taugen. In Übereinstimmung mit der Ablehnung einer strikten Unterscheidung zwischen analytisch und synthetisch verwirft der Holist die Ansicht, daß isolierten Termen überhaupt Bedeutungen zukommen – und dies zwar schon gar nicht in Form von intensionalen Entitäten. Das erkennende Subjekt zum Gegenstand der Psychologie als einem Teil der Gesamtwissenschaft deklassierend bestreitet er, daß die Introspektion irgend einen methodisch privilegierten Zugang zu den privaten Bewußtseinszuständen gewährt. Wenn wir im Anschluß an den Behaviorismus die Forschung auf empirisch zugängliche Objekte einschränken, brauchen wir nicht die Existenz von mentalen Vorgängen vorauszusetzen. Denn die Theorie, die das öffentlich beobachtbare Verhalten am besten erklärt, erfordert kein Quantifizieren über nicht-physikalische Entitäten. Darin zeigt sich der Naturalist radikaler als der ontologisch eher neutrale Sprachanalytiker, der bloß beansprucht, Sätze mit mentalistischen Ausdrücken ohne wesentlichen Verlust in Sätze mit Prädikaten für Verhaltensdispositionen übersetzt zu haben. Abgesehen von der heiklen Frage nach der effektiven Übersetzbarkeit[8] zieht es der Materialist ohnehin vor, ontologisch verdächtige Objekte grundsätzlich zu eliminieren, wodurch er zugleich auch das psycho-physische Problem aufzulösen glaubt.

Wie ist es um die tatsächliche Durchführung eines derartigen Programms bestellt? Betrachten wir zunächst die Version,[9] nach welcher – allgemein formuliert – Ereignisse im Bewußtsein mit physiologischen Vergängen im Gehirn identisch sind. Sie entspringt offenbar der Absicht, dem ontologischen Dualismus zu entgehen, ohne sich in die notorischen Schwierigkeiten der Übersetzungsthese zu verstricken. Was aber ist genau mit einer solchen Identitätserklärung gemeint? Im Satz ‚Schmerzen sind Gehirnzustände' kann das Wort ‚sind' jedenfalls nicht die Relation der sog. strikten Identität beinhalten, weil dann das Gesetz ‚$(x)\ (y)\ [(x = y) \rightarrow (F)\ (Fx \longleftrightarrow Fy)]$' uns zu unsinnigen Aussagen der Art verpflichten würde, daß Gehirnzustände heftig, beißend, dumpf oder verblassend sein können, die offensichtlich einen sog. Kategorienfehler involvieren. Da es auch nicht als ‚ist synonym mit' oder als ‚ist extensional gleich mit'[10] gedeutet werden kann, hege ich den Verdacht, daß derartigen Sätzen letztlich nicht mehr Sinn als die doch eher triviale Behauptung einer invarianten Korrelation abzugewinnen ist. Abgesehen von der essentialistisch anmutenden Frage ‚Was ist Schmerz?', an deren Beantwortung vermutlich dem Identitätstheorektiker besonders gelegen ist, erlaubt uns im übrigen die Feststellung einer kontingenten Wechselbeziehung genau dieselben Voraussagen zu machen wie eine Identitätstheorie. Wie ich befürchte, ermangeln jedoch beide aus noch zu erläuternden Gründen einer effektiven Erklärungskraft.

Trotz seiner sorgfältigen Ausführungen scheint mir auch Hilary Putnams Deutungsversuch des ‚ist' als desjenigen der empirischen Reduktion[11] letzlich unverständlich. Vor allem bezweifle ich, daß die Belange der Ontologie durch sie in klarer Weise geordnet werden. Denn anders als im Fall von ‚Der Morgenstern ist der Abendstern' wo die beiden Namen[12] den gleichen Himmelskörper bezeichnen, wird die Frage nach der Referenz, sofern das ‚ist' der Reduktion eine symmetrische Relation ausdrückt, nicht eindeutig geregelt. Wir können nämlich ohne Willkür nicht entscheiden, ob nun die qualitativen (psychischen) Erlebnisse, die wir mit neurobiologischen Vorgängen gleichsetzen, oder die neurobiologischen (physikalischen) Vorgänge, die wir mit qualitativen Erlebnissen gleichsetzen, wirklich existieren. Jedenfalls ist es wegen ihres völlig verschiedenen Status – die einen sind privat, die anderen öffentlich beobachtbar – nicht möglich anzunehmen, daß es sich um eine einzige Art von Entitäten handelt. Wir können es mit anderen Worten nicht mit einem Fall zu tun haben, wo wir mittels verschiedener Terme über dasselbe sprechen.

Mit der etwas anders gelagerten reduktionistischen Strategie des sog. eliminativen Materialismus' lassen sich die bedenklichsten der angedeuteten Unstimmigkeiten vermeiden. In der typischen Ausprägung etwa von Quines physikalischem Realismus wird als Argument ins Feld geführt, daß allein eine naturwissenschaftliche, d.h. eine physiologische oder neurobiologische Theorie die Phänomene, die wir gewöhnlich als geistige bezeichnen, wirklich zu erklären vermag und daß deshalb die Ontologie nicht mit mentalen Entitäten belastet zu werden braucht. Mit einem holistischen Standpunkt gekoppelt läuft das radikale Verfahren darauf hinaus, die Existenz von qualitativen Erlebnissen überhaupt zu leugnen. Seine Vorteile sind augenfällig: die Theoriebildung wird vereinfacht und das Problem der mentalen Verursachung verschwindet zugleich mit den Empfindungen als wirklichen Dingen. Trotzdem scheint mir intuitiv eine solche Lösung nicht völlig befriedigend, weil ich nicht glaube, daß wir in der Praxis ganz ohne eine vorgängige mentalistische Theorie auskommen.

Hilary Putnam, der zögert, endgültig Stellung zu beziehen, hält andererseits das Ansinnen, ausschließlich menschlichen oder tierischen Organismen mentale Zustände zuzuschreiben, für zu spezifisch. Von den Verhaltenssimulationen in den Computerwissenschaften angeregt, strebt er eine extreme Verallgemeinerung an. Seiner funktionalistischen Hypothese gemäß, die er für empirisch prüfbar zu halten scheint, wären psychologische Eigenschaften mit funktionalen Eigenschaften identisch.[13] Dem Versuch einer rein physikalistischen Reduktion hält er entgegen, daß das Hirn Züge aufweist, die nicht in biochemischen Termen beschreibbar sind. Als wesentlich habe er die Eigenschaft hervor, daß es ein von seiner stofflichen Realisierung unabhängiges Programm besitzt. Putnam gehört zu den subtilsten und kenntnisreichsten Autoren unserer Zeit; entsprechend stützt er seine These mit ausführlichen Argumenten, deren Spannweite von Gedankenexperimenten über sprachphilosophische Überlegungen bis

zur Erörterung von neurophysiologischen Befunden reicht. Wie allzu häufig im Rahmen der zeitgenössischen analytischen Philosophie wird man sich allerdings fragen, was der Aufwand an technischer Virtuosität nützt, wenn das gesamte Konzept dermaßen gegen die Intuition verstößt, daß seine Plausibilität von vorneherein als gering erscheint. Wie soll man sich im Ernst zu Putnams Äußerung stellen, wonach Roboter ein Verhalten zeigen könnten, das Ärger verrät? Es ist kaum anzunehmen, daß das Gerät in der Lage ist, ein Gefühl des Unmuts zu empfinden, das demjenigen ähnlich wäre, das ich verpüre, wenn ich vergeblich versuche, mich von der Richtigkeit einer solchen Behauptung zu überzeugen. Man möchte zunächst annehmen, daß jemand, der so spricht, einfach das Wort ‚Ärger' falsch verwendet. In einem zu weit gefaßten Kontext ist es jedoch heikel, eine Grenze zwischen dem rein Linguistischen und dem Empirischen zu ziehen. Es kann deshalb auch kein allgemeines Kriterium geben, das uns erlauben würde, *kontextunabhängig* zu entscheiden, wann jemand eine Sachlage falsch beurteilt und wann er sich in der Wahl der Terminologie vergreift. Der Sprachgebrauch ist wandelbar; er wird u.a. durch die Ergebnisse der wissenschaftlichen Forschung in beträchtlichem Maße beeinflußt. Dieser Sachverhalt läßt sich aber, wie ich fürchte, holistisch nicht erfassen. Um ihn in den Griff zu bekommen, sind wir auf konventionelle Praktiken angewiesen, die eine scharfe Trennung der jeweiligen Kontexte ermöglichen. Wenn demnach Hilary Putnam behauptet, die Physiker hätten „entdeckt", daß Temperatur in Wirklichkeit durchschnittliche Molekularenergie ist, so kann das nicht besagen, daß wir innerhalb eines monistischen Gesamtsystems endgültig Individuen einer Gattung durch andere ersetzt haben. Es muß vielmehr bedeuten, daß wir in einem historischen Kontext mit Rücksicht auf momentane wissenschaftliche Bedürfnisse eine frühere Perspektive mit der ihr eigentümlichen qualitativen Sprache durch ein neues Rahmenwerk abgelöst haben, das uns erlaubt, eine bessere Theorie zu formulieren. Ohne Berücksichtigung der besonderen geschichtlichen und theoretischen Umstände, die zur Adoptierung der betreffenden Theorie geführt haben, bliebe ein Satz wie ‚Temperatur ist durchschnittliche Molekularenergie' unverständlich.

Meiner Überzeugung nach, die übrigens im gegenwärtigen Stand der Diskussion eine Bestätigung findet, gibt es keine empirischen Belege für die „Wahrheit" der einen oder anderen philosophischen Doktrin. Die Kontroverse dreht sich nämlich gar nicht um eine Sachfrage, sondern um praktische Entscheide, die im wesentlichen die Wahl von geeigneten Sprachsystemen und anderen für die Wissenschaft erforderlichen Instrumenten betreffen. Statt sich vergeblich abzumühen, die Wahrheit oder Falschheit von Aussagen wie ‚Bewußtseinszustände sind *F*' nachzuweisen, sollten wir uns deshalb eher nach methodischen Gründen umsehen, die für oder gegen eine solche Redeweise sprechen. Zugunsten meiner Auffassung, nach welcher es keine objektiven, von Konventionen unabhängigen Kriterien für die faktische Gleichheit von Wahrnehmungszuständen mit Vorgängen im Gehirn gibt, plädiert im übrigen die im allgemeinen zugestandene beobachtungsmäßige Ununterscheidbarkeit der verschiedenen Identitäts- oder Reduktionstheorien.

Wenn also Sätze der Form ‚Was man früher ein *F* genannt hat, ist nichts anderes als ein *G*' einen Sinn haben sollen, müssen sie anders analysiert werden, als das bisher geschehen ist. Obschon auch ich der Meinung bin, daß der wissenschaftliche Fortschritt das Ausscheiden gewisser Terme aus der Sprache und die Eliminierung entsprechender Arten von Entitäten aus der Ontologie erforderlich macht, glaube ich nicht, daß das aufgrund der üblichen holistischen Sicht möglich ist. Denn, hätten wir es, wie z.B. Quine das idealerweise voraussetzt, mit einer einzigen, umfassenden Sprache/Theorie zu tun, so müßten vor der Reduktion in ihr beides, mentale und physikalistische Prädikate vorkommen, wodurch wir unweigerlich in die erörterten Schwierigkeiten zurückgeworfen würden. Um all den mißlichen Folgen zu entgehen, sehe ich keinen anderen Ausweg, als für eine Methodologie zu optieren, die dank der Ermöglichung einer strikten Unterscheidung zwischen analytisch und synthetisch sowie zwischen Sprache und Theorie die Bahn zu einem echten Pluralismus freigibt.

‚Was man früher ein *F* genannt hat, ist in Wirklichkeit ein *G*'. Illustrieren wir die allgemeine Formel durch die Einsetzungsinstanz ‚Was man im Mittelalter Besessenheit durch einen Dämon genannt hat, ist in Wirklichkeit eine Nervenkrankheit', so wird deutlich, daß mit der-

artigen Sätzen nicht die Behauptung einer Identität intendiert ist. Das Beispiel besagt nicht, daß Dämonen und Geisteskrankheiten identisch seien, sondern bringt in elliptischer Form etwa das Folgende zum Ausdruck: Im Mittelalter haben sich die Leute mit (wissenschaftlich leeren) Erklärungen vermittelst von bösen Geistern begnügt, während wir heute ungewöhnliches Verhalten mit Hilfe von (etwas besseren) psychiatrischen Theorien erklären. Ähnlich verhält es sich im Falle des Schmerzes und der Reizung von C-Fasern, indem wir uns im Alltag mit einer noch vagen, mentalistischen Sprache zufriedengeben[14], die im Kontext psychologischer oder physiologischer Forschung versagen würde. Diesen Umstand sollten wir wenigstens so weit in Betracht ziehen, daß der Philosophe sich trotz materialistischer Neigungen nicht dazu versteigt, die Existenz von Entitäten mentaler Art absolut zu leugnen. Es würde sich sicher nicht lohnen, über die ontologischen Sünden zu meditieren, die man in diesem Zusammenhang Ockhams Grundsatz gemäß begeht. Eine Übersetzung in die kanonische Notation würde nichtsdestoweniger ans Licht bringen, daß wir in einem solchen Kontext über mentale Entitäten quantifizieren.

Genau das haben wir jedoch zu vermeiden, wenn es darum geht, eine empirische prüfbare Theorie zu entwerfen. In diesem Fall müssen wir eine reglementierte Sprache mit exakt identifizierbaren Individuen verwenden. Im engeren Kontext der Wissenschaft findet somit eine physikalistische Strategie eine gewisse Berechtigung, was aber nicht heißt, daß sie um der Einheitswissenschaft willen zum tyrannischen Prinzip erhoben werden sollte. Die Vereinheitlichung unserer Theorien ist zwar durchaus wünschbar, kann jedoch als Fernziel nur schrittweise verwirklicht werden. Wie weit entfernt wir uns noch von einem solchen Ideal befinden, ließe sich ironischerweise trotz der ungeheuren Fortschritte, die auf dem Gebiet der Neurophysiologie in den letzten Jahrzehnten erzielt wurden, gerade am Stande unseres Wissens über Hirnvorgänge veranschaulichen.

Ein sinnvoller Reduktionsprozeß kann also nach meiner Auslegung nicht darin bestehen, daß wir innerhalb eines Gesamtsystems Objekte aufgrund einer empirisch festgestellten Identität eliminieren. Er muß vielmehr in der Weise verstanden werden, daß wir im Kontext einer Theoriebildung aus pragmatischen Gründen ein Sprachsystem entwickeln, in welchem die Prädikate so gewählt werden, daß die in der Ontologie als wirklich vorausgesetzten Entitäten tatsächlich der Beobachtung zugänglich sind. An einer Stelle scheint auch Richard Rorty zu einem sehr ähnlichen Ergebnis zu gelangen, ohne allerdings den nächsten Schritt zu einem folgerichtigen Pluralismus zu vollziehen: „Die Unterscheidung zwischen Beobachtungstermen und Nicht-Beobachtungstermen ist relativ zur jeweiligen Sprachpraxis (die sich ändern kann, wenn die Forschung Fortschritte macht), anstatt ein für allemal abgesteckt werden zu können, indem man zwischen den ‚gefundenen' und den ‚gemachten' Elementen in unserer Erfahrung unterscheidet".[15] Die im Zitat angedeutete Änderung der Praxis wird, wie ich meine, nicht durch Korrekturen an einem Totalsystem vorgenommen, sondern dadurch, daß wir eine Vielzahl von eigenständigen Teilsystemen entwerfen, die zueinander in der Relation der Unverträglichkeit, der Inkommensurabilität, der empirischen Äquivalenz etc. stehen können. Deshalb scheint es mir auch unangebracht, die Frage nach der Wahrheit einer Aussage wie ‚Früher hat man das Wort ‚Atom' falsch verstanden, weil man damit die Vorstellung der Unteilbarkeit verband', als durch einen Hinweis auf die tatsächliche Teilbarkeit der Atome für endgültig erledigt zu betrachten. Angesichts der unvermeidlichen Kontextabhängigkeit, die ein derart absolutistisches Ansinnen vereitelt, muß man sich mit einer Deutung bescheiden, die nicht mehr besagt, als daß die Quantenphysik gewisse Naturphänomene besser, d.h. exakter und vollständiger erklärt als etwa die Theorie Demokrits oder Epikurs. In analoger Weise wäre die Behauptung, daß *die* Sprache durch wissenschaftliche Entdeckungen ständig verändert wird, zu behandeln. Sofern sie das Vorhandensein einer einzigen, universell verwendeten Sprache nahelegt, ist sie jedenfalls irreführend, weil wir in Wirklichkeit eine Vielfalt von gegeneinander abgrenzbaren, z.T. künstlichen linguistischen Instrumenten entfalten, die jene Entdeckungen erst möglich machen.

Die soweit dargelegten Überlegungen, scheinen mir nun den methodologischen Schluß nahezulegen, daß es sich beim gegenwärtigen Stand unseres Wissens nicht empfiehlt, in einem

Sprachsystem zugleich mentale und physiologische Prädikate einzuführen. Man könnte zwar für den bloßen Zweck einer Abkürzung den Term ‚Schmerz' vermittelst einer expliziten Definition:

$$x \text{ ist ein Schmerz} =_{\text{Def}} x \text{ ist ein neurobiologischer Zustand der Art } \ldots$$

so einführen, daß nur ein neuer Ausdruck vorgelegt wird. Da wir trotz der Gleichheit der Lautzeichen dadurch nicht die früheren Entitäten mitpostulieren würden, hätten wir es jedoch in diesem Fall mit einem rein physikalistischen Gebrauch des Wortes zu tun.

Zur Rechtfertigung meines Vorschlags möchte ich weniger auf die kongenitalen Schwächen der Identitätstheorien als auf die betrübliche Tatsache hinweisen, daß für eine absehbare Zukunft vermutlich keine Theorie in Aussicht steht, die eine im streng wissenschaftlichen Sinne gesetzesartige Beziehung zwischen mentalen und physiologischen Prozessen herstellen würde. Solange aber eine solche aussteht, erscheinen mir philosophische Spekulationen in Richtung eines universellen Materialismus eher als müßig – und dies zwar umsomehr, als sie – soweit ich das abzuschätzen vermag – selbst nicht viel zur Verwirklichung einer experimentell prüfbaren physikalistischen Theorie des Bewußtseins beitragen. Ähnlich beurteilt auch Reinhard Werth die Situation, [16] indem er den geringen wissenschaftlichen Wert der diesbezüglichen Arbeiten in der analytischen Philosophie beklagt und ausdrücklich die Ersetzung der Alltagssprache durch eine Fachsprache verlangt, in der Terme wie ‚geistig', ‚das Selbst', ‚bewußt', ‚privat' etc. so präzisierbar sind, daß sie effektiv in Experimenten verwendet werden können. Im Zeichen eines wohlverstandenen Empirismus neige ich dazu, ihm beizustimmen, wenn er darauf beharrt, daß es für die Erklärung der Leib-Seele-Beziehung wesentlich darauf ankommt, die neurophysiologischen Abläufe zu erforschen, die in Organismen (nicht in technisch realisierten Systemen!)[17] für die Verarbeitung von Reizen und für die Steuerung des Verhaltens verantwortlich sind. Da ich andererseits weder den heuristischen Wert spezifisch philosophischer Überlegungen noch die praktische Bedeutung nicht-naturwissenschaftlicher Kontexte allzu gering ansetzen möchte, glaube ich, daß eine liberale Gesinnung, wie sie meine offene Methodologie fordert, letztlich zu einer fruchtbareren Gesamtsicht führt als die monistischen oder fundamentalistischen Rivalen. Indem sie bei fortschreitender Entwicklung nur einen dem jeweils vorliegenden Problem angemessenen Grad von Exaktheit verlangt, fördert sie die Einsicht, daß es für den Anfang wichtiger ist, mit einer plausiblen, wenn auch vagen, Hypothese zu beginnen, als von vorneherein auf einer strengen Formalisierung zu beharren, die uns in die Irre leiten wird.

Um Mißverständnissen vorzubeugen, füge ich abschließend hinzu, daß ich mit der Ansicht, wonach die abwechslungsweise Verwendung einer mentalistischen oder physikalistischen Sprache verschiedene Kontexte erzeugt, nicht jegliche Beziehung zwischen Bewußtsein und Vorgängen im zentralen Nervensystem abstreiten will. Auf der Metastufe läßt sich durchaus feststellen, daß mit dem Satz ‚Ich spüre Schmerz' gleichzeitig ein entsprechender Satz über die Reizung von C-Fasern etc. wahr wird. Eine bloße Registrierung eines kontingenten Zusammentreffens erklärt jedoch nichts. Eine interessante Korrelation im Sinne etwa einer kausalen Verknüpfung würde erst eine (vermutlich äußerst komplexe) die beiden Gebiete verbindende Einheitstheorie liefern, die – wie ich leider befürchte – noch in weiter Ferne liegt.

ANMERKUNGEN

[1] Vgl. dazu meine Aufsätze „Logik als Instrument des Philosophen", in: *Wissenschaftliche Akten der Schweizerischen Naturforschenden Gesellschaft* (1982) „Ansätze zu einer zeitgenössischen Transzen dentalphilosophie", in: *Kant a due secoli dalla „Criteria"*, Editrice La Scuola (Brescia 1984), S. 113–123, und „Method in Philosophy and Logic" *Dialectica* , 36, 4, 1982.

[2] Vgl. u.a. „Probleme der Ontologie", in: *Zeitschrift für allgemeine Wissenschaftstheorie*, IX, 1(1978); und „Ansätze zu einer zeitgenössischen Transzendentalphilosophie".

[3] Wie Regeln angewandt werden und wer befugt ist, sie einzuführen, sind Fragen, die im wesentlichen die Wissenschaftssoziologie angehen.

[4] Ich bin mir bewußt, daß eine nähere Bestimmung dieses Terms eher heikel ist. Ich verwende ihn nicht mit der (ebenfalls vagen!) Bedeutung einer möglichen Welt im Sinne der zeitgenössischen Semantik. Für mich sind Kontexte keine Alternativen zu anderen Welten, sondern konkrete Situationen mit bestimmten, bezüglich des vorliegenden Unternehmens relevanten Zügen, die in der wirklichen Welt auftreten. Es wird vermutlich nicht leicht sein, präzise Kriterien für die Individuierung von derartigen Gebilden anzugeben. Neben konventionellen Überlegungen dürfte es im allgemeinen genügen, die verwendete Sprache zu beschreiben sowie den besonderen Gebrauch, den man von ihr macht, zu charakterisieren, um sich vor allem auch darüber Klarheit zu verschaffen, welche Arten von Objekten im Wertbereich der Variablen vorkommen.

[5] Es ist klar, daß sich nicht alle Sprachen in das System der Prädikatenlogik übersetzen lassen; diejenigen, für die das nicht möglich ist, dürften jedoch solche sein, für die ontologische Fragen ohnehin von geringem Interesse bleiben.

[6] Vgl. „Probleme der Ontologie" und „Ontologie im Lichte einer zeitgenössischen Analyse", in: *Conceptus* , Jg. XIV, Nr. 33 (1980).

[7] Für einen Versuch in dieser Richtung vgl. Alan Miller, „Where's the Use in Meaning?", *Dialectica*, 39 (1985), H. 1.

[8] Wie in der Tat müßte eine praktisch brauchbare, neutrale Sprache aussehen, in welcher Aussagen, die Kategorienfehler der weiter unten erörterten Art nach sich ziehen, so zu übersetzen wären, daß keine solchen mehr auftauchen? In „Sensations and Brain Processes",*Philosophical Review*, 68 (1959), S. 141–156, gibt J. J. C. Smart einen Satz wie ‚Ich sehe ein oranges Nachbild' mit der folgenden Formulierung wieder: ‚Es geht etwas vor, wie das, was vorgeht, wenn ich meine Augen geöffnet habe, wach bin und wenn sich eine Orange in gutem Licht vor mir befindet, das heißt, wenn ich wirklich eine Orange sehe'!!

[9] Für eine systematische Klassifizierung der Theorien über die Beziehung des Mentalen zum Physikalischen vgl. Donald Davidson, „Mentale Ereignisse", deutsche Übersetzung in:*Analytische Philosophie des Geistes*, hrsg. v. P. Bieri Meisenheim 1981, S. 78.

[10] Man spürt keine Hirnprozesse im Knie!

[11] „Die Natur mentaler Zustände" im erwähnten Sammelband.

[12] Die Venus ist kein Stern, sondern ein Planet; die Terme sind deshalb keine Kennzeichnungen, sondern Eigennamen, d.h. „rigid designators" im Sinne Kripkes.

[13] Vgl. etwa *Reason, Truth and History* (Cambridge University Press 1981), S. 79, wo er von seiner Theorie als von einer korrekten, *naturalistischen* Beschreibung der Leib-Seele-Beziehung spricht, sich aber im Anschluß an Nelson Goodman bereit erklärt, ebenfalls „korrekte" nicht-naturalistische, z.B. mentale Versionen einzuräumen. Ich bin verlockt, diese etwas sibyllinische Äußerung als eine Annäherung an meine eigene Position zu deuten.

[14] Wer sich gelegentlich einer Konsultation beim Arzt über Mängel an seinen Neuronen, Synapsen etc. beklagen wollte, wäre in der Tat schlecht beraten!

[15] „Leib-Seele Identität, Privatheit und Kategorien", in: *Analytische Philosophie des Geistes*, S. 106ff.

[16] Vgl. *Bewußtsein (Psychologische, neurobiologische und wissenschaftstheoretische Aspekte*) (Berlin-Heidelberg-New York-Tokio 1983).

[17] Mit einem Computer ließen sich keine empirisch sinnvollen Experimente durchführen, so daß das Wort ‚Bewußtsein' in diesem Zusammenhang keine einsichtige Bedeutung haben kann. Die Fähigkeit des Simulierens gewährt keinen Einblick in wirkliche psychische oder physiologische Vorgänge!

* * *

WIE IST NATURWISSENSCHAFT MÖGLICH? (GRUNDZÜGE EINES NATURALISTISCHEN IDEALISMUS)

Nicholas Rescher
Universität Pittsburgh

1. *Wie ist Physik möglich*?

Wie ist Naturwissenschaft – und im besonderen Physik – überhaupt möglich? Wie kommt es, daß es uns Menschen – Staubkörnchen nur im Weltall – gelingt, die Geheimnisse der Natur zu entschlüsseln und zu ihren Gesetzen verstandesmäßig Zugang zu finden? Und wie kommt es, daß unsere Mathematik – scheinbar eine freie Schöpfung menschlicher Erfindungskraft – benutzt werden kann, um den modus operandi der Natur mit geradezu unheimlicher Leistungsfähigkeit und Genauigkeit zu beschreiben? Wie kommt es, daß die gesetzmässige Ordnung der Natur für uns mit Begriffsmitteln erschließbar ist, die wir selbst erdacht haben?[1]

Wir haben es hier mit der großen Entdeckung des Pythagoras zu tun: daß „Dinge Zahlen" sind, daß die Verhältnisse und Verfahren unserer Umwelt (wie die harmonischen Intervalle im Bereich der Akustik) in mathematischer Sprache artikuliert werden können. Wir kehren also zurück zum Pythagoreischen Mysterium und zu den Betrachtungen in Platons „Timaios", die den Quellgrund einer Tradition von mathematisierender Kosmologie bilden, welche in kontinuierlichem Fluß von der Astrologie zur Astronomie führt – mit Kepler als einer bemerkenswerten Übergangserscheinung.

Es ist anzumerken, daß diese Frage der kognitiven Zugänglichkeit und mathematischen Faßbarkeit der Naturgesetze ihre beunruhigende Brisanz erst entfaltete, als die „scientific community" von der Gott-ist-tot-Idee beeinflußt wurde. Solange man die Welt als Produkt der schöpferischen Aktivität eines auf mathematischer Grundlage arbeitenden Verstandes betrachtete, als das Werk eines Schöpfers, der beim Entwerfen der Natur *more geometrico* verfährt, solange war die ganze Angelegenheit unproblematisch. Gott stattet die Natur mit mathematisch geordneten Strukturen aus und den menschlichen Verstand mit einer genau darauf abgestimmten Fähigkeit zu mathematischem Denken. Wie Natur und Verstand sich vereinbaren lassen, bereitet so keine Schwierigkeiten. Gott richtet es einfach in dieser Weise ein. Wenn aber *das* das Erklärungsprinzip für die Erfaßbarkeit der Naturgesetze ist, dann freilich verschwindet zugleich mit der Absage auf Gott auch unsere Erklärungsgrundlage für die Erkennbarkeit der Natur. Einige der größten Denker unserer Zeit halten diesen Verlust für endgültig.

Wissenschaftler und Philosophen allerersten Ranges bekräftigen heute oft und ohne Zögern, daß wir nicht hoffen können, das Rätsel um die mathematische Erkennbarkeit der Natur überhaupt zu lösen. Erwin Schroedinger charakterisiert den Umstand, daß der Mensch die Gesetze der Natur entdecken kann, als ein „Wunder, das möglicherweise jenseits des menschlichen Fassungsvermögens liegt"[2]. Eugene Wigner stellt fest, daß „die enorme Nützlichkeit der Mathematik in den Naturwissenschaften etwas ist, das ans Mysteriöse grenzt, und daß es keine rationale Erklärung dafür gibt"[3]; ja, mit einem unerwarteten Sinn fürs Lyrische steigert er sich zu der Behauptung: „Das Wunder, daß sich die mathematische Sprache zur Formulierung physikalischer Gesetze eignet, ist ein herrliches Geschenk, das wir weder verstehen noch verdienen."[4] Sogar Albert Einstein stand tief beeindruckt vor diesem Problem. In einem Brief, den er 1952 an Maurice Solovine, einen alten Freund seiner Berner Zeit, schrieb, heißt es: „Du findest es seltsam, daß ich die Erkennbarkeit der Welt – in dem Maße, in dem wir berechtigt sind, von solch einer Erkennbarkeit zu sprechen – als ein Wunder oder

ewiges Geheimnis betrachte. Nun, a priori sollte man erwarten, daß die Welt nur in dem Ausmaß Gesetzmäßigkeiten aufzustellen erlaubt, in dem wir mit unserem ordnenden Verstand Eingriffe vornehmen . . . (Aber) die Art der Ordnung, die beispielsweise die Gravitationstheorie Newtons errichtet, ist von gänzlich anderem Charakter. Auch wenn die Axiome der Theorie von Menschen gesetzt sind – der Erfolg des Unternehmens basiert auf einem hohen Grad an Regelhaftigkeit in den objektiven Weltverläufen, den wir a priori in keiner Weise zu erwarten berechtigt sind. Das ist das Wunder, das mit der Entwicklung unseres Wissens mehr und mehr zum Tragen kommt . . . Das Seltsame ist, daß wir uns damit zufrieden geben müssen, das ‚Wunder' anzuerkennen, ohne daß wir über eine legitime Methode verfügten, dahinter zu sehen . . ."[5] Gemäß all diesen Theoretikern stehen wir hier vor einem echten Geheimnis. Wir müssen akzeptieren, *daß* die Natur intelligibel ist, haben aber keine Aussicht, je verstehen zu können, *warum* dies so ist. Das Problem einer Erfaßbarkeit der Natur durch die mathematisch formulierten Theoriebildungen des Menschen gilt als unüberwindlich, als hoffnungslos unlösbar. Alle drei genannten, berühmten Nobelpreisträger für Physik verwenden ohne Scheu das Wort „Wunder" in diesem Zusammenhang.

Vielleicht ist aber die Frage geradezu illegitim und sollte überhaupt nicht gestellt werden. Vielleicht ist das Thema der Erkennbarkeit der Natur nur grundsätzlich jeder vernünftigen Behandlung unzugänglich, indem es in gewissem Sinn auf einer falschen Voraussetzung beruht. Denn nach einer Erklärung dafür zu suchen, warum naturwissenschaftliches Forschen Erfolg hat, setzt voraus, daß es einen Erklärungsgrund für diesen Sachverhalt tatsächlich gibt. Wenn aber dieser Umstand als reiner *Zufall* betrachtet werden muß, dann kann es kein solches Erklärungsprinzip geben. Eben diesem Gedankengang folgt Karl Popper, wenn er schreibt: „Doch selbst wenn wir annehmen (und das tue ich), unsere Erkenntnissuche sei bisher sehr erfolgreich gewesen, wir wüßten jetzt etwas über die Welt, dann ist dieser Erfolg unbegreiflich unwahrscheinlich und daher unerklärlich; die Berufung auf eine endlose Folge unwahrscheinlicher Zufälle ist keine Erklärung. (Das Beste, was wir tun können, scheint mir zu sein, die fast unglaubliche Entwicklungsgeschichte dieser Zufälle zu untersuchen . . .)."[6]

Aus dieser Sicht wird die Frage nach der Erkennbarkeit der Natur ein illegitimes Scheinproblem, eine verbotene Frucht, nach der zu greifen man vernünftigerweise nicht trachten sollte. Wir müssen uns einfach mit dieser Tatsache abfinden und einsehen, daß jeder Erklärungsversuch von vornherein zum Scheitern verurteilt ist.

Gegenüber der großen Frage, wie Naturwissenschaft überhaupt möglich ist, bekennen sich daher einige der scharfsinnigsten zeitgenössischen Wissenschaftler ratlos und gehen unverzüglich dazu über, das Problem in die Schleier des Wundersamen und Mystischen zu hüllen. Solch ein Vorgehen ist freilich sehr problematisch. Ungeachtet bedeutender Gegenstimmen ist die Frage nach der Erkennbarkeit der Natur allem Anschein nach nicht nur sinnvoll und bedeutsam, sondern auch eine solche, von der wir prinzipiell hoffen dürfen, sie mehr oder weniger vernünftig beantworten zu können. Die hier unternommene Behandlung der Frage jedenfalls beruht auf der Überzeugung, daß das Thema eine gehörige Portion Entmystifizierung nötig hat und auch verdient.

2. *Charakterisierung des Problems*

Wie kommt es, daß wir unser mathematisches Instrumentarium erfolgreich zur Beschreibung der Natur einsetzen können? Weshalb trifft Mathematik auf die Wirklichkeit zu? Der reine Logiker scheint eine Antwort parat zu haben. Er sagt: „Mathematik *muß* auf die Realität anwendbar sein. Mathematische Propositionen sind doch begriffliche Wahrheiten. Sie sind daher in dieser Welt wahr, weil sie in jeder möglichen Welt wahr sind." Aber diese Antwort geht an der Sache vorbei. Zugegeben, die wahren Sätze der reinen Mathematik sind in und von jeder möglichen Welt wahr. Aber sie sind dies nur, weil sie streng hypothetisch und ohne deskriptiven Gehalt sind. Gerade ihr Status als Begriffswahrheiten besagt, daß sie für unsere Zwecke nicht herangezogen werden können. Wir beschäftigen uns doch jetzt nicht mit der

apriorischen Wahrheit der reinen Mathematik und ihrem Leistungsvermögen, Vernunftwahrheiten zu formulieren. Worum es hier geht, ist vielmehr die Anwendbarkeit der Mathematik im Bereich der Erfahrungswirklichkeit, ihre zentrale Rolle bei der Aufgabe, den Rahmen für die kontingenten, aposteriorischen Wahrheiten über gesetzmäßige Abläufe abzustecken und so die Zusammenhänge der Natur für die Vernunft zugänglich zu machen.

Schließlich ist es ja vollkommen klar, daß die Gültigkeit der reinen Mathematik in einer Welt nicht bedeutet, daß die *Gesetze* dieser Welt mit relativ einfachen mathematischen Mitteln darstellbar sein sollen. Sie bedeutet nicht, die Verfahrensweisen der Natur müßten in stimmigem Bezug zur Mathematik stehen und in einfachen, klaren, eleganten und rational einsichtigen Formeln faßbar sein. Kurz, sie bedeutet nicht, die Welt müsse mathematisch diskutierbar oder „mathematophil" sein im Sinne einer Empfänglichkeit für eine Behandlung der eleganten Art, die ihr in den exakten Naturwissenschaften tatsächlich widerfährt. Wie also können wir die Tatsache begründen, daß die Welt auf die Mathematik paßt und umgekehrt?

Die Antwort erfordert einen etwas komplexen, doppelgleisigen Gedankengang. Der Umstand, daß X und Y sich in einem Zustand wechselseitiger Affinität und Übereinstimmung befinden (also daß Natur und Mathematik passend koordiniert sind), ist eine zweiseitige Angelegenheit, in der erwartungsgemäß beide Seiten von Belang sein werden.

Wenn wir die Natur verstehen wollen, wenn es zwischen diesen beiden Bezugspartnern Einklang und Übereinstimmung geben soll, dann müssen beide Teile, sowohl die Zusamenhänge der Natur als auch die mathematische Rationalität des Menschen, kooperativ zusammenwirken. Diesen Gedankengang müssen wir näher verfolgen.

3. Unsere Seite

Die Seite des Menschen in dieser zweiseitigen Angelegenheit ist relativ einfach zu beschreiben. Der Mensch ist in jedem Fall ein integraler Bestandteil der Natur, eingebunden in das System der Dinge als eine seiner inneren Komponenten. Daher ist der Typ von Mathematik, den er entwickelt, zwangsläufig anwendbar. Seine Erfahrung ist unvermeidlich Erfahrung *von Natur*. Genau *das* nämlich ist „Erfahrung": unsere durch unseren Erkenntnisapparat vermittelte Reaktion auf die als Stimuli fungierenden Umwelteinflüsse. So muß wohl die Art von Mathematik, die der Mensch im Hinblick auf diese Erfahrung entwickelt, eine solche sein, die auf die Natur, so wie er sie erfährt, anwendbar ist.

Die Mathematik einer in astronomischen Fernen lebenden Zivilisation dagegen, deren empirische Ausgangsbasis von unserer stark abweicht, könnte von der Mathematik, die wir kennen und lieben, fundamental verschieden sein. Der Umgang mit Quantität könnte dort gänzlich anumerisch erfolgen, z.B. rein vergleichend statt quantifizierend. Besonders wenn die äußere Umgebung dort nicht in genügendem Maße mit festen Gegenständen oder stabilen und damit meßbaren Strukturen ausgestattet ist – wenn die Bewohner dieser Welt etwa quallenähnliche Geschöpfe sind, die in einer Art „Bio-Suppe" herumschwimmen – so könnte die dortige „Geometrie" äußerst seltsam ausfallen, wohl vornehmlich auf Topologie konzentriert und eher auf flexible Strukturen eingerichtet als auf feste Größen oder Formen. Digitale Umsetzungen fehlten möglicherweise völlig, bestimmte Arten analogen Denkens andererseits könnten zu höchster Subtilität entwickelt sein. Wären ferner die extraterrestrischen Intelligenzen etwa diffuse Agglomerate von Einheiten, deren Zusammenfassung zu Ganzheiten auch Überlappungen zuließen, dann könnte ihr Denken so sehr von einer sozialen Begrifflichkeit dominiert werden, daß sie die Natur grundsätzlich unter sozialen Kategorien betrachteten, so daß jene Aggregate, die wir als *physikalische* Strukturen denken, bei ihnen unter einer *sozialen* Perspektive begriffen würden. Die Art von „Strukturen", die ihrer mathematischen Theoriebildung zugrunde läge, könnte *in der Tat grundlegend* von unserer eigenen abweichen.

Mathematik ist zwar die Theorie der Strukturen, aber sie muß sich doch als Theorie der vorstellbaren Strukturen entwickeln. Und „vorstellbar" bedeutet hier: vorstellbar für ein im

Laufe der Naturentwicklung entstandenes und in seine natürliche Umwelt eingebundenes Lebewesen. Mathematik ist freilich nicht Naturwissenschaft, sondern eine Theorie des hypothetisch Möglichen. Doch diese Möglichkeiten sind nichtsdestoweniger Möglichkeiten, wie sie *von uns* erfaßt werden, von Lebewesen also, die die Erfassung von Möglichkeiten mit einem evolutionär entwickelten, ihnen von Natur aus eingepflanzten Erkenntnisapparat erzeugen. Es darf uns also nicht überraschen, wenn die Art von Mathematik, die wir betreiben, eine solche ist, die sich zur menschenbegrifflichen Verarbeitung natürlichen Geschehens geeignet zeigt.

Unsere kognitiven Mechanismen – die Mathematik eingeschlossen – passen auf Natur, weil sie selbst ein Produkt natürlicher Abläufe sind, wobei diese ihrerseits vermittelt sind durch die kognitiven Prozesse intelligenter Lebewesen, die ihre Intelligenz dazu benutzen, ihre Interaktionen mit der Natur, deren Teil sie selbst sind, zu steuern. Unsere Mathematik ist dazu prädestiniert, auf die Natur einigermaßen abgestimmt zu sein, weil sie selbst ein Produkt der Natur ist; also weil sie in gewisser Hinsicht unsere eigene Stellung als integrale Konstituenten der Natur widerspiegelt.

Die kognitiven Mechanismen, die wir anwenden, wenn wir uns mit der Welt auseinandersetzen, wenn wir sensorische Interaktion in verstandesmäßig analysierbare Erfahrung umsetzen, haben neben vielen anderen Eigenschaften auch die, selbst ein Beitrag der Natur zur Anpassung eines ihr auf Gnaden oder Ungnaden ausgelieferten Lebewesens zu sein. Daß der Verstand des Menschen die Zusammenhänge der Natur begreift, ist nicht überraschender als die Tatsache, daß das Auge einen Teil der in der Natur vorkommenden Strahlung, oder daß der Magen Produkte der Natur als Nahrungsmittel verarbeiten kann. Dies alles haben wir dem Selektionsdruck zu verdanken.

Es könnte allerdings der Fall sein, daß unsere kognitive Anpassung nur innerhalb unserer unmittelbar nächstliegenden raum-zeitlich begrenzten Umgebung gut funktioniert, innerhalb der Mikro-Umgebung, die unsere besondere ökologische Nische definiert. So besteht die Möglichkeit, daß wir unter einem weiteren Blickwinkel gar nicht so gut dran sind, daß wir nur einen kleinen und sehr peripheren Teil eines großen, unzugänglichen Ganzen zu fassen bekommen. *Daher* kann in dem einseitigen Beitrag des Menschen nicht die *ganze* Antwort auf die Frage nach der Erkennbarkeit der Natur liegen. Denn selbst wenn wir, gemessen an unseren eigenen evolutionär entstandenen Bedürfnissen in unserer Naturerfassung ziemlich erfolgreich sind, so *könnte* diese in einem umfassenderen Zusammenhang doch immer noch sehr inadäquat sein. Es bleibt also noch zu zeigen, daß die Natur selbst unseren Erkenntnisleistungen die entsprechenden Ansatzpunkte bietet, daß sie der Vernunft prinzipiell – und nicht nur so irgendwie am Rande – zugänglich ist.

Um *diesen* Punkt zu klären, müssen wir daher dazu übergehen, den Beitrag der Natur in dem Wechselverhältnis von Natur und Geist in Sicht zu nehmen.

4. Die Seite der Natur

Wenn ein intelligentes Lebewesen, das seine natürliche Umgebung erforscht und dabei auf der Grundlage physischer Interaktion mit ihr mathematisch formulierte Entwürfe und Überzeugungen entwickelt, ein einigermaßen adäquates Verständnis von ihrer Funktionsweise erreichen soll, *dann* muß auch die Natur ihrerseits in gehörigem Umfang zur Zusammenarbeit beitragen. Offensichtlich muß sie zunächst die Entwicklung intelligenter Lebewesen zulassen. Darüber hinaus muß sie diesen Lebewesen eine Umwelt bieten, die über ausreichend stabile Konstellationen verfügt, um einen „Erfahrungs"-zusammenhang entstehen zu lassen und solche Lebewesen dazu zu befähigen, aus strukturierten Interaktionen geeignete *Informationen* zu gewinnen. Diese Überlegung hat weitreichende Implikationen. Im Hinblick auf unsere hier vorausgesetzte Zielsetzung muß gezeigt werden, daß die Mathematik nicht nur von einigem Nutzen ist, sondern daß sie für das Verständnis der Welt zwangsläufig eine äußerst wesentliche Voraussetzung darstellt. Wir müssen uns dessen versichern, daß die Anwendung von

Mathematik im Bereich rationaler Forschung ein adäquates und genaues Verständnis der Naturzusammenhänge liefern kann. Wir müssen also eingehender untersuchen, wie die Natur mit den Forschungsmethoden des Verstandes in einen Dialog tritt.

Der Beitrag der Natur selbst zur mathematisch artikulierbaren Naturerkenntnis muß in ihrer Strukturiertheit durch *relativ einfache* Gesetzmäßigkeiten bestehen, die in ihrer Regularität so leicht zu entschlüsseln sind, daß sie sogar von Lebewesen mit ziemlich bescheidenen Fähigkeiten und Mitteln gut verstanden werden können.

Wie aber kann man die These rechtfertigen, daß die Natur nach relativ durchschaubaren Gesetzmäßigkeiten geordnet sein *muß*? Lassen sich tatsächlich stichhaltige Gründe dafür anführen, weshalb die Welt, die wir mittels unseres vom mathematischen Denken geprägten Verstandes erforschen, auf der Basis relativ einfacher und der mathematischen Analyse von sich aus zugänglicher Prinzipien operieren sollte? Hier ist der zentrale Punkt wohl der folgende: Eigentlich darf es gar nicht erstaunen, daß der mathematisierende Verstand der Erfassung der Natur so wesentliche Erfolge erzielt, denn eine Welt, in der sich ein derartiger Erkenntnisapparat überhaupt entwickelt, ist eine Welt, die der Mathematik angepaßt ist. Dieser Schlußfolgerung liegt in groben Zügen etwa folgende Argumentation zugrunde: Eine Welt, in der sich durch *Evolutions*prozesse Intelligenz entwickelt, muß eine solche sein, in der diese Intelligenz den intelligenten Einwohner einen (zumindest bescheidenen) Evolutionsvorteil einbringen muß. Intelligenz muß für diejenigen, die darüber verfügen, ein „Selektionsplus" darstellen. Auch die niedrigsten Lebewesen, die nur höchst rudimentäre *Vorstufen* von Intelligenz besitzen (Schnecken und Algen zum Beispiel), müssen sich so verhalten, daß bestimmte Reiztypen (Muster sich wiederholender, definiert strukturierter Umwelteinflüsse) angemessen korrespondierende Reaktionstypen hervorrufen. Mit anderen Worten: solche Lebewesen müssen ein Reizmuster in ihrer Umwelt als solches isolieren und darauf in einer Weise reagieren können, die sie bei der Selektion begünstigt. Aber dies bedeutet, daß die Natur auf eine ganz bestimmte Art „kooperativ" ist, daß sie genügend gleichförmig, regelhaft und strukturiert sein muß, daß es auf Umweltereignisse „angemessene" und „lernbare" Reaktionen geben kann. Es muß also strukturierte Ereignismuster in der Natur geben, die schon die einfachsten Lebewesen in der Ausstattung ihres Umweltapparates und in ihrem modus operandi widerspiegeln können. Die Welt muß also problemlos lernbare Muster von Ereignisablauf-Strukturen beinhalten, das heißt relativ einfache Gesetzmäßigkeiten. Die Existenz solcher lernbaren Strukturiertheiten von natürlichen Vorkommnissen bedeutet, daß Mathematik, die letztlich die abstrakte und systematische Theorie der Strukturen im allgemeinen ist, *eine* nützliche Funktion haben muß. Und so können wir schließen, daß eine Welt, in der sich mittels evolutionärer Prozesse Intelligenz entwickeln kann, eine solche sein *muß*, die einem an mathematischen Begriffen orientierten Verständnis zugänglich ist.[7]

Eine Welt, in der evolutionäre Prozesse hochgradige Intelligenz hervorbringen, muß eine hochgradig intelligible Welt sein. Es liegt im Wesen der Evolution, daß eine Welt, in der mathematisierendes Denken überhaupt auftaucht, angesichts des damit verbundenen notwendigen Selektionsvorteils eine Welt sein muß, in der dieses Denken imstande ist, die Welt mathematisch zu verstehen. Die Existenz intelligenter Lebewesen mag für sich und als solche erstaunlich sein oder nicht; aber sobald es sie einmal gibt und sobald wir einsehen, daß es sie dank des Verlaufs der Evolution gibt, darf es nicht länger erstaunlich sein, daß ihren Bemühungen, die Welt in der Sprache der Mathematik zu beschreiben, ein substanzieller Erfolg beschieden ist. Eine Welt, in der Lebewesen mit hochgradiger Intelligenz entstehen können, muß eine sein, deren Gesetzesstrukturen eine leistungsfähige Beschreibung mit mathematischen Mitteln ermöglichen und erfordern.

Ein weiteres wichtiges Problem lauert noch im Hintergrund: *Wie gut* versteht unsere Naturwissenschaft die Welt wirklich? Kommt sie nicht nur zu mutmaßlichen, sondern auch zu echten Naturgesetzen? Das freilich ist eine Frage, die nicht als solche beantwortet werden kann. Denn wir haben keinerlei *direkten* Zugang zur Natur-an-sich, um diese mit der aus unseren Untersuchungen resultierenden mutmaßlichen Natur zu vergleichen – um unsere als solche behaupteten „Naturgesetze" mit ihren realen Gegenstücken zu konfrontieren. Alles was wir

tun können, ist, die Adäquatheit unserer Theoriebildung am Grad ihrer *Leistungsfähigkeit* zu überprüfen, an ihrer Eignung, uns Vorhersagen und kontrollierte Eingriffe zu erlauben. Und freilich sind wir unter diesen pragmatischen Hinsichten zu hinlänglich guten Leistungen prädestiniert. Jene mutmaßlichen Gesetzmäßigkeiten wären eben nicht Bestandteile unseres naturwissenschaftlichen Kanons, wenn sie uns nicht unter diesen Rücksichten in recht guter Weise dienlich wären. (Dies ist der Ort, wo die *rationale* Selektion ins Spiel kommt). Es ist effektiv unvermeidlich, daß wir den Eindruck gewinnen, unsere Wissenschaft bilde die Natur relativ adäquat ab. Jene Theorien, die nach dem als „naturwissenschaftliche Methode" bezeichneten Verfahren gewonnen und bestätigt werden, gehörten doch gar nicht zum Corpus unserer Naturwissenschaft, wenn sie bei uns nicht diesen Eindruck hinterließen. Wenn wir über den „Erfolg der mathematisch formulierten Naturbeschreibung" reden, so spielen wir mit gezinkten Karten. Denn (1) muß solch ein Erfolg im Sinne eines *wahrgenommenen* Erfolges konzipiert werden und (2) entstammt diese Erfolgswahrnehmung demselben (biologischen und kognitiven) Evolutionsprozeß, der auch der für uns spezifischen Art von Naturverständnis zugrunde liegt.

Der Erfolg der Mathematik bei der Beschreibung der Natur ist so in keiner Weise erstaunlich. Daß sich intelligente Lebewesen überhaupt entwickeln konnten, mag Verwunderung hervorrufen oder nicht. Aber daß sie, einmal im Laufe der Evolution in Erscheinung getreten, erfolgreich sein würden bei dem Vorhaben, die Natur in mathematischen Beziehungen zu erfassen, ist nur zu erwarten. Eine Welt, in der sich Intelligenz zeigt, deren Lebewesen zu ihrer Intelligenz durch evolutionäre Mechanismen gelangen, muß für ein Verständnis nach mathematischen Begriffen wesentlich offen sein.

Man darf dabei nicht übersehen, daß der hier vorgeführte Argumentationsgang von strikt hypothetischem Charakter ist. Es wird nicht (kategorisch) behauptet, daß die Welt, um eine elegante mathematische Darstellung ihrer Verfahrensweisen zu ermöglichen, einfach genug sein muß und damit punktum. Es wird vielmehr die streng konditionale These behauptet: *Wenn* intelligente Lebewesen im Laufe *evolutionären* Geschehens in der Welt auftauchen, *dann* muß die Welt mathematophil sein. Entscheidend ist dabei die Einsicht, daß diese konditionale Form der These doch für unsere Zielsetzung völlig ausreicht; denn wir haben es mit der Antwort auf die Frage zu tun, weshalb wir Menschen, wir intelligenten Lebewesen, die Welt in der Sprache unserer Mathematik verstehen können sollten. Und die genannte konditionale Form genügt, um diese Aufgabe zu erfüllen.

5. Synthesis

Lassen Sie mich nun kurz innehalten, um die einzelnen Stücke zusammenzufügen. Die allgemeine Frage nach der Erkennbarkeit der Natur hat zwei Richtungssinne:

I. Warum ist der Verstand so gut auf die Natur abgestimmt?

II. Warum ist die Natur so gut auf den Verstand abgestimmt?

In der Erörterung dieser Fragen wurde die Ansicht vertreten, daß ihre Beantwortung gar nicht so schwierig ist – zumindest auf der Ebene einer schematischen Charakterisierung. Der Verstand muß an die Natur angepaßt sein, weil er ein Produkt evolutionärer Abläufe ist. Und die Natur muß durch den Verstand erfaßbar sein, weil dies allererst die Ausbildung von Verstand überhaupt ermöglicht.

Wenn Natur und Verstand zu weit ab von jeder Übereinstimmung wären, wenn der menschliche Erkenntnisapparat zu „unintelligent" wäre, um die vielschichtigen Zusammenhänge der Natur zu begreifen, oder die Natur zu komplex für die Fähigkeiten des Verstandes, dann könnten beide eben nicht auf Takt kommen. Es wäre wie der Versuch, Shakespeare in ein Pidgin-Englisch mit einem Vokabular von 500 Wörtern umzuschreiben – zu vieles und zu wesentliches ginge da verloren. Es wäre wie der Versuch, ein System mit zehn relevanten

Freiheitsgraden in einem Modell darzustellen, das nur über vier von diesen verfügen kann. Wenn es sich wirklich so verhielte wie in diesen Fällen, könnte der menschliche Erkenntnisapparat seine evolutionäre Mission nicht erfüllen.

Die Lösung unseres Problems wurzelt also in der Kombination zweier Überlegungen: (1) Eine Welt, die im Zuge der Evolution eine relativ potente Form von Intelligenz auftreten läßt, muß relativ regelmäßig und einfach strukturiert, das heißt mathematophil sein. (2) Wie eine hinlänglich potente Form von Intelligenz eine derartige Welt mittels mathematisch-naturwissenschaftlicher Theoriebildungen begreifen kann, ist dementsprechend durch die Tatsache zu erklären, daß im Hinblick auf die Evolution Verstand und verstandesmäßige Erfaßbarkeit in signifikanter Weise koordiniert sein müssen.

7. Ein Restpunkt

Diese Überlegungen bringen uns aber auf ein wesentliches restliches Element im Problemfeld um die Erkennbarkeit der Natur. Denn sie lassen die Schlüsselfrage noch unberührt: Warum haben sich in der Natur überhaupt verstandesbegabte Wesen entwickelt? Wie kommt es, daß die Welt der Schauplatz forschender Intelligenz ist?

Dies ist eine noch problembeladenere Frage als die soeben angesprochene – und in der Tat eine sehr alte. In Tabelle I findet sich eine einigermaßen vollständige Liste der verschiedenen Antworten, die dazu vorgeschlagen wurden.

TABELLE I

Weshalb hat sich in der Natur überhaupt Intelligenz entwickelt?

I. Die positivistische Antwort:
Die Frage ist illegitim, unangemessen, sinnlos.

II. Die skeptizistische Antwort:
Die Frage ist sinnvoll, aber wesentlich unbeantwortbar.

III. Die abblockende Antwort:
Es gibt diesen (verflixten) Grund einfach nicht. Natur ist eben so. Die Natur der Natur hat keinen explanatorischen Hintergrund, sie bildet das Ende der Erklärungsreihe.

IV. Die theologische Antwort:
Gott hat es so eingerichtet.

V. Die Teleologische Antwort:
Die Gründe für das Auftreten von Intelligenz sind wesentlich teleologisch: Das Vorhandensein von Intelligenz ermöglicht es der Natur, ein bestimmtes Ziel zu verwirklichen.

VI. Die evolutionstheoretische Antwort:
Mit der Entstehung von Leben in der Welt ist auch die Entwicklung von Intelligenz zu erwarten, denn Intelligenz erhöht die Überlebenschancen. Sie bietet eine wirksame Lösung für das Problem, eine biologische Art entlang der Zeitachse voranzubringen.

Die teleologische Antwort klingt besonders verführerisch. Sie schreibt der Natur eine bestimmte Neigung zur Ausbildung von Geistigkeit zu, einen inneren Antrieb zur Ermöglichung von Erkenntnisvollzügen. Sie stützt sich auf die in der Natur aufweisbare (oder ihr zuweisbare) Grundtendenz zur Selbstreproduktion (*conatus se replicandi* ähnlich dem *conatus*

se praeservandi, der Selbsterhaltung). Die Natur, als ein Ganzes genommen, kann dies freilich nicht bewerkstelligen – es sei denn durch die Hervorbringung von Lebewesen, die Gedankenmodelle, nicht physische, aber begriffliche Abbildungen von ihr schaffen. So kommt die Natur dazu, Lebewesen hervorzubringen, die, wie Leibniz' Monaden, die Fähigkeit besitzen, die Welt in einem gewissen Grade verstandesmäßig „widerzuspiegeln". Die Welt verkörpert demnach eine von sich aus auf Geist hin abzielende Bewegung. Die Natur hat – in bester Tradition hegelschen Denkens – ein inneres Bestreben, sich selbst als Gedanke zu verwirklichen.

Die wahre Erklärung – die evolutionstheoretische Antwort – ist weitaus prosaischer, nüchterner und unromantischer. Sie beginnt mit der Feststellung, daß es viele verschiedene Wege gibt, mit der Natur sein Auskommen zu finden, daß den biologischen Arten diverse Wege geboten sind, um sich in der Natur zu behaupten: Erzeugung zahlreicher Nachkommenschaft, Widerstandsfähigkeit, Anpassungsfähigkeit, Isolation, et cetera. Unter diesen Marschrouten erweist sich aber eine als besonders aussichtsreich: diejenige, die darauf ausgeht, Intelligenz auszubilden, sich dadurch anzupassen, daß Gehirnkapazitäten statt Muskelkapazitäten oder erhöhter Reproduktionsfähigkeiten zum Einsatz kommen. Wie schon Meister Darwin selbst betonte, hat die Art, die Zusammenhänge ihrer Umgebung am besten erfaßt, einen Selektionsvorteil: Die Art, die ihren Weg durch den Darwinschen Konkurrenzkampf nicht durch reine Zähigkeit macht, nicht nur mit Klauen und Zähnen, sondern mit Intelligenz, findet eine vielversprechende ökologische Nische. Sobald sich das *Leben* entwikkelt und unter der Schutzherrschaft der fruchtbaren Mutter Natur darangeht, verschiedene Fortpflanzungsstrategien auszukundschaften, ist es nur natürlich, daß sich Intelligenz entwikkelt. Kurz und gut: Intelligenz tritt auf, weil sie für die Lösung des Überlebensproblems ein wirksames Instrument darstellt. Intelligenz entwickelt sich nicht deshalb, weil die Natur dies begünstigt, sondern weil Intelligenz das Überleben derer, die sie besitzen, innerhalb der Natur begünstigt (auf alle Fälle bis zu einem bestimmten Punkt).

Schluß

Wir wollen noch einmal rasch den Verlauf der Diskussion durchgehen und im Rückblick die drei Hauptpunkte Revue passieren lassen.

1) Intelligenz entwickelt sich in belebter Natur, weil sie Lebewesen eine gute Methode bietet, mit ihrer Umwelt zurechtzukommen.

2) Sobald intelligente Lebewesen entstehen, haben ihre Erkenntnisleistungen aller Wahrscheinlichkeit nach einen bestimmten Grad von Adäquatheit, da sie durch den Selektionsdruck auf die Arbeitsweise der Natur hin ausgerichtet sind.

3) Es sollte nicht überraschen, daß diese Anpassung schließlich eine äußerst leistungsfähige, mathematisch formulierte Physik entstehen läßt, denn die grundlegenden Strukturen natürlicher Vorgänge, die es auch zur Ausbildung von Intelligenz kommen lassen, *müssen* relativ einfach sein.

Im weiten Umfeld dieser Thesen mag in der Tat allerhand geheimnisvoll bleiben. (Fragen wie: „Warum hat sich *Leben* in der Welt entwickelt?" und – noch grundsätzlicher –: „Warum existiert die Welt überhaupt?" könnten als Kandidaten vorgebracht werden.) Sei das, wie es will, die gegenwärtig behandelte Thematik, weshalb die Natur dem Menschen erkennbar ist und weshalb sich solche Erkennbarkeit in einer mathematisch formulierten Physik niederschlagen sollte, erfüllt doch die Qualifikationen eines großen Mysteriums, geschweige denn eines Wunders, sicherlich nicht.

Gewiß, die hier vertretene Konzeption ist äußerst schematisch entworfen und erfordert umfangreiche Füllarbeiten. Es muß eine lange und verwickelte Geschichte über die physikalische und die kognitive Evolution erzählt werden, um die Konzeption mit Einzelheiten anzureichern. Aber es gibt sicherlich allen Grund, darauf zu hoffen und zu warten, daß eine solche Geschichte schließlich erzählt werden kann. Und das ist hier der entscheidende Punkt. Selbst wenn jemand an den Umrissen der hier skizzierten evolutionstheoretischen Darstellung, an

der Linienführung im einzelnen, seine Zweifel hat – so bleibt doch die Tatsache, daß *irgendeine Darstellung dieser Art* in bezug auf die Erfaßbarkeit der Natur mit mathematischen Mitteln eine hervorragend ausbaufähige Erklärung liefert. Die Tatsache allein, daß so eine Konzeption im Prinzip möglich ist, zeigt, daß man die Aussichten auf eine Beantwortung unserer Frage nicht so schwarz in schwarz malen muß, wie es die Zonen undurchdringlicher Geheimnisse sind.

Es besteht einfach keine Notwendigkeit, Einstein, Schroedinger und Co. Folge zu leisten, wenn sie die Erkennbarkeit der Natur als Wunder oder Geheimnis betrachten, das alles menschliche Verstehen übersteigt. Wenn wir bereit sind, von der Naturwissenschaft selbst zu lernen, wie die Natur funktioniert und wie der Mensch bei der Auswertung seiner Forschungsergebnisse vorgeht, dann könnten wir zunehmend das Problem, weshalb ein Lebewesen dieser Art eine physikalische Umgebung solcher Art effektiv erforschen kann, vom Schleier des Geheimnisses befreien. Wir wären schließlich imstande, es als völlig plausibel und ganz allen Erwartungen entsprechend zu finden, daß Wissenschaft betreibende Lebewesen auftreten und sich eine Position verschaffen, die es ihnen erlaubt, daraus allerhand Kapital zu schlagen. Bei der Suche nach einem Leitfaden zur Beantwortung der Frage, wie Naturwissenschaft möglich ist, können wir daher die *Wissenschaft selbst* in Blick nehmen. Und es gibt keinen guten Grund, weshalb sie uns in dieser Hinsicht im Stich lassen sollte.

ANMERKUNGEN

[1] Dieses ausgesprochen kantische Thema wird hier in einer ausgesprochen unkantischen Weise behandelt. Denn die vorliegenden Überlegungen beschäftigen sich nicht à la Kant mit bestimmten apriorischen Prinzipien, die der Physik vermutlich zugrunde liegen. Wir betrachten vielmehr die faktischen (aposteriorischen) Prinzipien, aus denen die Physik *besteht* – die Naturgesetze selbst.

[2] Schroedinger, Erwin, *What is Life?* (Cambridge 1945), S. 31.

[3] Wigner, Eugene P., „The Unreasonable Effectiveness of Mathematics in the Natural Sciences", *Communications on Pure and Applied Mathematics*, Bd. 13, (1960), S. 1–14. (vgl. S. 2)

[4] Ebd. S. 14.

[5] Einstein, A., *Lettres à Maurice Solovine* (Paris 1956), 114f.

[6] Popper, K. R., *Objektive Erkenntnis*, (dtsch. Übersetz. Hamburg 1973), S. 40f.

[7] Gespräche mit Gerald Massey haben in hilfreicher Weise zur Verbesserung und Klärung dieses Argumentationsteils beigetragen.

[8] Es ist freilich nicht die *biologische* Evolution im Sinne Darwins, die zur Mathematik führt, sondern die, sich an die biologische anschließende, kognitive Evolution, bei der die rationale Selektion der intellektuellen Mittel im Sinne der Lamarque'schen Theorie verläuft. Vgl. N. Rescher, *Methodological Pragmatism* (Oxford 1973).

* * *

GIBT ES FÜR DAS LEIB-SEELE-PROBLEM EINE „LÖSUNG“?

Dieter Birnbacher
Universität Essen

Ist das Leib-Seele-Problem ein echtes oder ein Scheinproblem? Bedarf die Frage nach dem Zusammenhang zwischen Leib und Seele der *Lösung* oder der *Auflösung?*

Ich bin überzeugt, daß das Leib-Seele-Problem *kein* bloßes Scheinproblem ist, das zu seiner Lösung lediglich der Klärung seiner sprachlichen Struktur bedarf. Aber ich bin ebenso davon überzeugt, daß wir gut daran täten, nicht schnurstracks auf die „Lösung“ des Leib-Seele-Problems zuzugehen, sondern zunächst in reflektiver, metaphilosophischer Einstellung nach den *Kriterien* zu fragen, die wir dafür gelten lassen wollen, eine vorgeschlagene „Lösung“ als Lösung zu akzeptieren. Erst wenn wir uns dieser Kriterien vergewissert haben – wenn wir wissen, was wir suchen, was wir von einer „Lösung“ erwarten –, wird sich entscheiden lassen, ob es eine „Lösung“ im Prinzip geben kann, ob sich bereits einer der gegenwärtig diskutierten Lösungsansätze als „Lösung“ auszeichnen läßt, oder ob es dazu weitergehender Fortschritte des empirischen oder philosophischen Wissens bedarf.

Der *erste* Teil meines Beitrags wird einen Vorschlag zu den Kriterien machen, dem eine potentielle Lösung genügen sollte. Der *zweite* Teil wird dann prüfen, welche der gegenwärtig diskutierten Lösungsansätze diese Kriterien am ehesten erfüllen. Das Interessante an dieser Vorgehensweise ist, daß sich der Kreis der Bewerber mit ernsthaften Chancen als außerordentlich klein erweist. Der nächstgelegene Weg zur „Lösung“, so zeigt sich, führt in der Tat über die „Grammatik“ – wenn auch nicht die der Alltagssprache, sondern die der Philosophie, genauer: der *induktiven Metaphysik*. Unter einer *induktiven Metaphysik* soll dabei eine Hypothese (oder Hypothesensystem) verstanden werden, die die Existenz erfahrungstranszendenter Entitäten, Eigenschaften oder Relationen zum Zweck der Erklärung der Erfahrungswirklichkeit postuliert und von größerer Reichweite ist als die allgemeinsten Hypothesen und Hypothesensysteme der Wissenschaften. Ob zwischen wissenschaftlichen und metaphysischen Hypothesen in jedem Fall eine scharfe Grenze zu ziehen ist, kann dabei zunächst offenbleiben.

Die im folgenden skizzierten Kriterien sind Bewertungskriterien, die mehr oder weniger gute Gründe für die Akzeptierung einer metaphysischen Hypothese darstellen sollen, keine Ausschlußkriterien (constraints). Im Prinzip soll die Nichterfüllung eines Kriteriums also jederzeit durch die Erfüllung anderer Kriterien aufgewogen werden können. Die Reihenfolge der Nennung soll dem relativen Gewicht entsprechen: 1. *Intelligibilität,* 2. *Explizitheit der logischen Struktur,* 3. *Logische Konsistenz,* 4. *Kompatibilität mit weithin akzeptierten realwissenschaftlichen Theorien,* 5. *Komparativer Problemlösungsbeitrag,* 6. *Integrationsleistung,* 7. *Sparsamkeit,* 8. *Kohärenz mit anderen akzeptierten metaphysischen Theorien,* 9. *Einfachheit,* 10. *Heuristische Fruchtbarkeit.* Als „Lösung“ ließe sich dann diejenige Hypothese apostrophieren, die diese Kriterien *bestmöglich* erfüllt, als „vorläufige Lösung“ diejenige, die sie *auf dem gegenwärtigen Kenntnisstand bestmöglich* erfüllt.

Diese Kriterien erlauben zunächst die Eliminierung von zwei Lösungsvorschlägen: erstens Varianten eines *reduktiven Materialismus,* die alle psychologischen Prädikate – einschließlich von „raw feels“-Prädikaten – als (wie immer komplexe) physikalische konstruiert; zweitens interaktionistische Konzeptionen, die wie die von Hans Jonas Energieaustauschprozesse zwischen Psychischem und Physischem postulieren, sich dabei jedoch über die Erhaltungssätze der Physik hinwegsetzen zu können glauben.

Attraktiv aufgrund der angegebenen Kriterien ist eine Identitätstheorie, die ausgehend von der empirisch bestätigten Annahme einer durchgängigen Korrelation von psychischen Ereig-

nissen mit Gehirnereignissen eine kontingente Identität von psychischen (P-) und bestimmten materiellen (M-) Ereignissen postuliert und einen Monismus der Entitäten (es gibt nur materielle Objekte) mit einem Dualismus der Eigenschaften (es gibt außer materiellen auch psychische Eigenschaften) verbindet. Sie übertrifft ihre dualistischen Konkurrenten sowohl hinsichtlich der Integrationsleistung (Kriterium 6) als auch hinsichtlich der Sparsamkeit (Kriterium 7). Interessanterweise erweisen sich die der Identitätstheorie am häufigsten angekreideten Schwächen (wie daß sie unseren Sprachgewohnheiten zuwiderläuft, indem sie psychische Ereignisse räumlich lokalisiert) nach dem angegebenen Kriterienkatalog als von allenfalls marginaler Bedeutung.

Problematisch an der Identitätstheorie ist allerdings die Einbuße an Problemlösungskapazität durch die Tendenz, die Frage nach der zwischen den miteinander identifizierten Phänomenen bestehenden *kausalen Beziehung* zu verschütten. Die Identität von P- und M- Ereignissen ändert ja nichts daran, daß diese Ereignisse zwei miteinander korrelierte Aspekte haben, für die sich die Frage stellt, ob M_i Ursache von P_i, P_i Ursache von M_i oder P_i koinzidierende Symptome eines zugrundeliegenden X_i sind. Die Identitätstheorie ist jedoch mit allen kausalen Deutungen verträglich, selbst noch mit der Annahme einer prästabilierten Harmonie. Sie hat insofern für das Problem des *Zusammenhangs* zwischen Geist und Materie keinen Lösungsvorschlag anzubieten. Sie bedarf der Erweiterung und Präzisierung.

Eine dieser möglichen Erweiterungen und Präzisierungen kann aufgrund des Kriterienkatalogs von vornherein als ceteris paribus weniger akzeptabel ausgeschlossen werden, da sie im Widerspruch zum ontologischen Sparsamkeitsprinzip steht: die Annahme eines unbekannten Dritten, dessen „Symptome" Geist und Gehirn sind. Betroffen davon sind sowohl Ansätze, nach denen uns stets nur die Symptomklassen M und P zugänglich sind – wie bei Spinoza, sofern man ihn als Vertreter einer solchen Theorie interpretiert –, als auch Ansätze, die eine zukünftige Emergenz „höherer" Symptomstufen nicht ausschließen – wie die von „angels" oder „finite gods" bei Samuel Alexander.

Unter den verbleibenden Ansätzen, dem Interaktionismus, der Kausalität in beiden, und dem Epiphänomenalismus, der Kausalität ausschließlich in der Richtung M–P annimmt, ist der Epiphänomenalismus nach den angegebenen Kriterien eindeutig zu präferieren. Erstens bringt der Epiphänomenalismus das, was wir wissen, auf den kleinstmöglichen metaphysischen Nenner (Kriterium 7). Zweitens vermeidet der Epiphänomenalismus ein gravierendes erkenntnistheoretisches Problem, mit dem der Interaktionismus konfrontiert ist: daß unter Bedingungen durchgängiger P-M-Korrelation eine Annahme kausal aktiver P-Ereignisse, wie sie für den Interaktionismus kennzeichnend ist, absolut willkürlich wäre.

Denn angenommen, der psychische Ereignisaspekt P_i sei die kausal hinreichende Bedingung für das materielle Ereignis M_j. P_i korreliere seinerseits mit dem materiellen Ereignis M_i. Dann gibt es für M_j entweder eine rein materielle kausal hinreichende Bedingung oder nicht. Angenommen, es gibt sie, so wäre diese keine andere als M_i, das mit P_i korrelierende M-Ereignis. Beide, P_i und M_i wären kausal hinreichende Bedingungen für M_j. Unter diesen Bedingungen gäbe es jedoch keinerlei Möglichkeit zu überprüfen, ob P_i der Annahme entsprechend tatsächlich eine kausal hinreichende Bedingung für M_j ist. Dazu müßten wir überprüfen können, daß P_i M_j auch ohne das Eintreten von M_i bewirkt, was unmöglich ist, da P_i und M_i korrelieren. Nehmen wir andererseits an, daß es für M_j *keine* materielle kausal hinreichende Bedingung gibt, so würden wir auch P_i nicht als kausal hinreichende Bedingung von M_j betrachten können, da andernfalls auch M_i eine kausal hinreichende Bedingung für M_j wäre, was der Voraussetzung widerspricht. Was dem Interaktionismus Schwierigkeiten macht, ist also weniger eine etwaige kausale Geschlossenheit der physikalischen Welt – im Gegensatz zur kausal nicht geschlossenen Welt des Psychischen – als vielmehr die Annahme einer durchgängigen Korrelation von psychischen mit materiellen Ereignissen. Auch wenn sich der notorisch interaktionistisch denkende common sense als unbelehrbar erweisen sollte – der identitätstheoretische Epiphänomenalismus erweist sich als derjenige Lösungsansatz, der den angegebenen Kriterien maximal gerecht wird und insofern, im Sinne eines historisch relativen Optimums, als die „vorläufige Lösung" des Leib-Seele-Problems betrachtet werden sollte.

Die These, daß der Epiphänomenalismus seinerseits keine Probleme neu aufwirft, ist mit mehreren Argumenten bestritten worden:

1. Es laufe der Idee kausaler Abhängigkeit zuwider, wenn das Abhängige nicht seinerseits wieder Ursache sei. – Aber weder besteht eine solche Notwendigkeit, noch wird sie durch das Kausalitätsprinzip ausgesagt. Auch die Vertreter des Interaktionismus nehmen nicht an, daß *alle* psychischen Wirkungen materieller Ursachen ihrerseits kausal aktive Rollen übernehmen.

2. Eine Kausalität von P nach M sei introspektiv „axiomatisch" gewiß. – Aber die Introspektion ist die falsche Instanz, wenn es darum geht, was „in ihrem Rücken" geschieht. Die Tatsache, daß wir gelegentlich die Vorstellung einer unmittelbaren Verknüpfung zwischen Willensakt und Verhalten haben, schließt nicht aus, daß der bewußte Willensimpuls nur vermöge seiner Korrelation mit einem Gehirnvorgang auf den Körper einwirkt.

3. Hätte der Epiphänomenalismus recht, wäre damit unsere Selbstsicht als Personen, die handelnd in die materielle Welt eingreifen, als Illusion desavouiert. – Aber dieser Einwand greift nur, wenn ein Begriff der Person vorausgesetzt wird, der die Person von der Existenz des menschlichen Gehirns abkoppelt. Die Vorstellung einer *körperunabhängigen* Einwirkug des Geistes auf das Gehirn wäre in der Tat eine Illusion – eine Illusion mit einem gewissen survival value allerdings, da es sein könnte, daß sie das Ich, indem sie ihm suggeriert, es schiebe, statt geschoben zu werden, in seinem Selbstbehauptungswillen stärkt.

4. Der Epiphänomenalismus sei mit evolutionistischen Ansätzen zur Erklärung des Bewußtseins unvereinbar. Nach diesem – namentlich von Popper – erhobenen Einwand soll es nur dann für eine nach dem Prinzip der natürlichen Selektion operierende Evolution einen Grund gegeben haben können, Lebewesen mit Bewußtsein hervorzubringen, wenn der Besitz von Bewußtsein eine kausal notwendige *und* hinreichende Bedingung für bestimmte überlebensdienliche Verhaltensweisen gewesen ist. Andernfalls wäre die „unaufwendigere" Alternative rein materieller Entwicklungslinien bevorzugt worden. Aber Poppers Einwand hat zwei entscheidende Schwächen: Erstens die Voraussetzung, daß die Evolution, um intelligentes Verhalten hervorzubringen, des Psychischen als eigenständiger, kausal aktiver Realität bedurfte; eine Natur, die teils lange vor der Dämmerung des Bewußtseins die Zelle, die geschlechtliche Fortpflanzung und die staatenbildenden Insekten „erfunden" hat, hätte auch das intelligente Verhalten „erfinden" können, ohne den Umweg über ein von der materiellen Basis abgekoppeltes Psychisches zu nehmen. Zweitens die Voraussetzung, daß die Natur eine Wahl hatte – eine Wahl zwischen Lebewesen *ohne*, und anderen, ebenso leistungsfähigen, *mit* der Zutat des Bewußtseins. Diese Wahl hatte sie dem Epiphänomenalismus zufolge jedoch gerade nicht: Falls die Evolution das intelligente Gehirn wollte, mußte sie nomologisch notwendig auch das Bewußtsein wollen. Hätte sie eine Wahl gehabt, hätte es in der Tat keinen Grund gegeben, weshalb sie die aufwendigere unter den zwei Möglichkeiten gewählt hätte. Der Epiphänomenalismus hält alles bereit, was sich die Evolutionstheorie nur wünschen kann, um zu erklären, warum die Evolution Lebewesen mit Bewußtsein bevorzugt hat, falls diese einmal durch Zufallsmutation entstanden sind. Freilich: das Faktum der nomologischen Verknüpfung von Materie und Geist selbst bleibt einer biologischen wie einer induktiv-metaphysischen Erklärung notwendig entzogen. Sie ist – jedenfalls bislang noch – Sache eines rein spekulativen Emergenzdenkens. Mit *dieser* Unzulänglichkeit unseres Denkens in Frieden zu leben, ist jedoch *jedem* ernstzunehmenden Versuch einer „Lösung" des Leib-Seele-Problems aufgegeben. Sie ist keine Hypothek, die allein auf dem Epiphänomenalismus lastet.

LITERATUR

Alexander, S., *Space, Time, and Deity,* Bd. 2 (London ²1927).

* * *

RAMSEY'S THEORY OF MIND IN RELATION TO HIS VIEW ON "THEORIES"

Ulrich Majer
Universität Hannover

In Ramsey's "Last Papers", posthumously published, there is a short article, entitled "Philosophy", that contains a sophisticated argument for the necessity of "self-consciousness" in philosophy.[1] Strange as this argument may seem, I want to show that you cannot understand "theories" and "theoretical terms" correctly according to Ramsey without reflecting on "our own mental states", that is, without taking into accont self-consciousness as part of a subjectivistic theory of mind.

In what follows, I shall first give a reconstruction of the argument, which leads—according to Ramsey—from some simple premises about the nature of philosophy to the conclusion, that "self-consciousness (is) inevitable in philosophy".[2] In a second step I want to show in greater detail, *why* you cannot understand theories and theoretical terms, in Ramsey's view, without the kind of self-consciousness just mentioned. In a third step I shall give an outline of Ramsey's theory of mind by drawing out some of the consequences of Ramsey's view on theories and theoretical terms with respect to an adequate theory of mind including self-consciousness. In a last step I return to the main conclusion Ramsey draws from his theory of mind, namely, I quote: " we cannot make our philosophy into an ordered progress to a goal, but have to take our problems as a whole and jump to a simultaneous solution".[3] (The problems in the interpretation of quantum mechanics are an outstanding example for this anti-fundamentalistic conclusion.)

I. The reconstruction of the argument, why self-consciousness is inevitable in Philosophy

Clearly, the argument depends primarily on what we understand by "philosophy". According to Ramsey it is the main and only task of philosophy to provide definitions of terms and sentences we use in science (as well as in every day life). This is so, simply because philosophy has as its main question "What do we mean by x?" A question, the answer to which is best given in the form of a definition of x (suppose, that x is a term). Therefore we can state as the first premise of the argument:

(1) "A philosophy is a system of definitions"[4]

You have to take this, admittedly, rather strange thesis quite seriously and literally, that is, you have to take it that

(i) philosophy is a system of definitions just as logic is a system of tautologies or mathematics is a system of identities.[5]

(ii) different philosophies differ from one another in providing different systems of definitions as different "logics" differ from one another in providing different systems of "tautologies".

Naturally, the next question that comes to mind in regard to such a definitive programme, is: "What do we mean by definitions?" (and hence by philosophy). Yet, in trying to answer this question we encounter a serious difficulty; a difficulty which involves the danger of a per-

petual "petitio principii" or something like that: We cannot answer this question simply by giving a definition of "definition". First we have to think about what definitions essentially are, and then to spell out the answer. If you think about the question "What definitions essentially are", you will sooner or later come to see that not all terms of a genuine language[6] can be defined, at least not in the obvious manner of giving a *nominal* definition by performing a speech act of the kind "x means such and such".[7] The very idea of a definition prohibits that every term of a language should be definable, because in the definiens you need terms, that already have a meaning, either by definition, in which case the argument can be repeated, or without any kind of nominal definition. Hence, in every genuine language there must be undefined terms or primitives, the meanings of which have to be explicated in a non-definitional way, or else you will go around in a circle. Let us call such non-definitional explications of the meaning of undefined terms for short "explanations"; then we can state the second premise of Ramsey's argument as follows:

(2) Philosophy, as a system of definitions, contains, besides nominal definitions, undefined terms, the meanings of which can only be *explained* but not explicitly defined.

Thus having saved philosophy from a flagrant circularity by admitting undefined terms in the systems of definitions we go further and ask: 'How do we explain—and grasp thereby—the meaning of undefined terms?' But at this point, philosophers often make an unnoticed presupposition, namely, that there is only *one type* of undefined terms in every system of definitions and hence only need for one type of explanation. This presupposition however is not always correct, or, according to Ramsey, it is plainly false, because there are at least two quite different types of undefined terms in every genuine language (and hence in philosophy):

(i) terms, the meaning of which we are already acquainted with, like terms for sense impressions or common sense objects; these terms stand in no real need of explanation, because we could easily point to their visible instances.

(ii) terms like "mass" and "energy", that look as if they should be defined in primitive terms—but of which a rigorous meaning analysis can show that it is not only unnecessary but quite *useless* to define them. These terms stand in serious need of explanation and cannot be defined nominally, because their definitions would be cumbersome and completely impractical. This type of expressions is what Ramsey calls "*variable hypotheticals*" and "*theoretical terms*".

Ramsey's philosophy mainly differs from others like that of Carnap and Russell in acknowledging that there are (at least) two types of undefined terms: primitives and theoretical terms (of which only the second stand in real need of explanation). Therefore we can state as the third premise of Ramsey's argument:

(3) Philosophy as a system of definitions, contains besides primitives so-called "theoretical terms" which have to be explained and should not be nominally defined.

It is clear that first this third premise contains the clue to the argument, because the conclusion that self-consciousness is inevitable in philosophy, only follows if we ask, how we explain (or grasp) the meaning of theoretical terms and variable hypotheticals.

Ramsey's answer to that question implicitly involves the conclusion that self-consciousness is inevitable in philosophy, because—I quote—"in this explanation we are forced to look not only at the objects, which we are talking about, but at our own mental states".[8] To understand Ramsey's answer correctly, we have to find the reason why in the explanation of theoretical terms and variable hypotheticals "we cannot neglect the epistemic or subjective side".[8] In order to do this, we have to study the relation between Ramsey's view on theories and the kind of self-consciousness he demands.

II. The relation between Ramsey's view on theories and the demand for self-consciousness in philosophy

Ramsey's answer to the question 'How do we explain the meaning of theoretical terms and variable hypotheticals?' can be split into two parts: a more negative one already contained in thesis (3) and a positive one which gives the essence of Ramsey's later philosophy.

(i) The negative answer is best understood, if you realize that general propositions are *no propositions at all* but something like causal laws.

The reason for this is simply that general propositions do not only state facts, that is something which is definitely true or false, but also express general hypotheses about future situations, which are not properly true or false, but only more or less believable inference licenses for possible situations. To be precise, this means that general propositions cannot in principle be defined as finite or infinite conjunctions of propositions proper, but have to be explained as variable hypotheticals. What that means I will explore in the next part, for the moment I only want to stress that the negative answer implies that general propositions are not truth functions of atomic propositions—and nothing like that either, as in Tarski's definition of truth. The solution which Ramsey proposes for the problem of understanding general propositions is much more radical than Tarski's.

Passing over to the positive part of the answer I have to mention that theories, on Ramsey's view, are nothing but a special kind of rather complex variable hypotheticals, namely hypotheticals in which we introduce, for the sake of simplicitly and clearness, fictitious entities like "mass" and "force", that will be designated by the theoretical terms of the theory. That these should not be defined but have to be explained means that we can and should give only an *indefinite description* of the meaning of theoretical terms in terms of sense impressions. To give such indefinite descriptions of theoretical terms is, according to Ramsey, logically equivalent to making an *existential claim* of the form: Theoretical entities of such and such a kind exist—but not uniquely! In other words, Ramsey's philosophy is, in contrast to that of Russell and Carnap, a kind of anti-reductionist philosophy, insofar as Ramsey suggests that theoretical terms have their own *irreducible* meaning, which cannot be expressed uniquely and completely in terms of sense impressions, but have a fictitious character. Every attempt to define that fictitious character in terms of sense impressions like Carnap's and Russell's, makes the introduction of theoretical terms quite pointless![9] Out of what material—if not sense impressions—this fictitious character is built, and how we grasp this irreducible part of the meaning of theoretical terms and variable hypotheticals by reflection on our own mental states, is the concern of the next point.

(ii) Again, the positive answer is best understood, if you consider why the genuine meaning of general propositions like 'All men are mortal' cannot be reduced to a finite or infinite conjunction of atomic propositions, but has to be explained by the way in which we *use* these general propositions or sentences respectively.

The reason is simply that their meaning is mainly *potential* (in contrast to actual) and hence cannot be expanded into facts or propositions about facts, but has to be elucidated in a completely unconventional manner, in which a kind of reflection on our own mental states plays an essential role. Take the above example and ask in what its meaning consists. One point is pretty clear: Its meaning is not exhausted by 'Socrates is mortal and Plato is mortal and Aristotle is mortal, and so on and so on' and by no conjunction of death reports, however long it may be, but only if we take into account all potential cases of men and says: 'If x is a man, then x is mortal', that means, he will die. But this is no proposition at all about facts, no statement true or false, but the expression of a conviction, namely the belief that all men will die, whoever they may be. This is in fact how we use this general proposition in every day life

as well as in science: We use it not to make a statement about facts—and also not about 'general' facts, whatever that may be—but to express our belief that whatever a man will do for his immortality, as a man, he will die.

The difference between statement and belief is important here, because a belief, and especially a general belief, is not in the same way true or false as a statement about facts is true or false. In the latter case the correspondence theory of truth seems adequate, in the former surely not, because a belief about potential situations cannot be compared with facts, at least not in the case of general beliefs such as 'All men are mortal'. In this case the question is not so much one of truth and falsehood (as in the case of statements) but of *credibility*: Shall we trust in the belief of another speaker? Is his general proposition a good 'variable hypothesis' for managing the future or not? Time will show, and if not, we will have to find a better one. This is the pragmatic notion of truth as "what everyone will believe in the end", and it is clearly a coherence—and not a correspondence notion of truth. This is not the time to go deeper into this matter, but let me mention two points before I return to the main question, why self-consciousness is inevitable in philosophy.

(α) The semantical analysis which Ramsey undertakes in regard to general propositions, presents them as a kind of 'cognitive attitude' in regard to the future, namely as inference licenses from the past to the future in which we believe. The generality of the inference is indicated as the variability of the range, and the epistemic status of belief is represented by the hypothetical form of the inference.

(β) The semantic distinction Ramsey makes between ordinary and general propositions is, logically speaking, rather different from the analogous distinction between propositional and predicate logic. While in the latter case the semantics of general propositions is constructed in parallel to that of singular propositions, Ramsey considers general propositions as instances of 'cognitive attitudes' in regard to propositions. This means that they require their own logic and semantics, just as epistemic or modal propositions require their own logic and semantics! In this respect Ramsey agrees with the intuitionists, especially with H. Weyl, who also sees the main gulf in logic between propositional logic on the one side and general as well as existential propositions on the other, the transcendental side of knowledge.[10]

Coming back to our main question, I hope it is evident by now, why self-consciousness is inevitable in philosophy at least in the case of general propositions: If our belief in regard to potential, and especially to future situations is not only part but the very essence of the meaning of general propositions, then "we are forced to look not only at the objects which are talking about, but at our own mental states".[8] We are forced to look at our own mental states, that is, our beliefs, wishes, and desires, for short, all our mental habits because they lie at the bottom of the meaning of general propositions, which as variable hypotheticals go far beyond what we actually know. This reflection on our beliefs is inevitable, if we want to *know* what we mean by a general proposition, insofar as no actual states of affairs can tell us that, but only potential ones; and these exist only in our mind and nowhere else. Sure, we could object that the potential situations are similar or analogous to the actual ones, but this is only to rename the problem of belief, because we, by our beliefs, have built the potential situations similar to the actual ones. Hence, the habit of belief is the material out of which we form the meaning of general propositions. Every philosophy, dealing with the meaning analysis of general propositions, has therefore to reflect upon these beliefs. That we can should do this is itself a hypothesis, called "self-consciousness", that is yet not fully understood.

What has been shown for general propositions could equally well be shown for theories as instances of variable hypotheticals with "theoretical terms". Interesting as this would be, I have to pass over this here, because of Ramsey's highly technical representation of "Theories". Let me only mention that Ramsey regards a theory of mind as a theory in his sense, that is, as a variable hypothetical about our own as well as other minds with certain theoretical terms like "self-consciousness". This brings me to my third topic.

III. Ramsey's theory of mind involving "self-consciousness".

In order to explain Ramsey's theory of mind I have to go back to a possible objection against Ramsey's view that self-consciousness is inevitable in philosophy. The objection runs roughly as follows:

The explanation of the meaning of general propositions by reference to our own mental states is *circular* insofar as it tries to explain the meaning of general propositions by the notion of variable hypotheticals, that is, roughly, by our belief in causal laws, which themselves stand in serious need of explanation. Instead, we should do our philosophical analysis—so the objection runs—in a given frame of logic and try to explain causal laws by general propositions, thinking all the time of the facts without reflecting on our own mental states.

That is, by the way, what is usually done up to this day. We try to explain causal laws by the logic of general propositions, thinking all the time about facts and the inductive generalizations without taking into account our beliefs in potential situations as the source of inductive reasoning. That is to say, we take it for granted that there is inductive generalization, which leads entirely without self-consciousness from facts to general propositions and in this way to causal laws.

Yet, according to Ramsey—and, I admit, also in my own view—this leads to an *impasse*, because, as we now all know, in this way we cannot understand causal laws and their intended meaning. What we have to do instead is, roughly speaking, the following: We have to become clear about how we think in the realm of *deductive* reasoning about the future, namely by stating variable hypotheticals and theories in general. Now according to Ramsey, this has two quite different yet intimately connected aspects: (i) On the one hand he contends that "over and above the regularity of succession the world, or rather that part of it with which we are acquainted, exhibits no feature called causal necessity". (ii) Yet, on the other hand "we make sentences, called causal laws, from which we proceed to actions and propositions connected with them". Insofar as these two aspects exhaust the matter, they entail a consequence which is highly relevant to every theory of mind, because the second aspect is, in Ramsey's view, nothing but "a regular feature of our conduct, a part of the general regularity of things; as always there is nothing beyond the regularity to be called causality, but we can again make a variable hypothetical about this conduct of ours".[11]
To my understanding this means at least two or three things:

(α) The objective meaning of every causal law is nothing but the regularity of successions of events in the world.
(β) This does not prevent that we—as a part of the world—make variable hypotheticals, i.e. causal laws, which go far beyond the actual world we know by projecting potential situations and fictitous entities. This conduct of ours is in fact itself a part of the regularity of the world.
(γ) Therefore, we can again make variable hypotheticals about this feature of our mind and call them "theories of mind". But we have to be aware that these theories of mind are themselves nothing but variable hypotheticals about a supposed regularity of our mind.

The last point is the crucial one, because it implies that every theory of mind has to contain a certain minimum of self-consciousness: It must reflect that it is itself only a theory—a theory about our mind—that contains an irreducibly hypothetical element.[12]

This brings me back to the thesis, quotet in the beginning, that "we cannot make our philosophy into an ordered progress to a goal, but have to take our problems as a whole and jump to a simultaneous solution".[3] The demand of hypothetical-deductive thinking instead of inductive reasoning has itself "something of the nature of a hypothesis";[13] it implies that we have to take into account self-consciousness as an essential part of the meaning analysis of theoretical terms and variable hypotheticals—and hence of a theory of mind. But of course, no one can

prove this in the logical sense of proof.–Sure, you could try to explain the meaning of theoretical terms without self-consciousness, yet we believe in hypothetical-deductive thinking. This has to be shown useful–more useful than any other hypothesis–in the discourse of practical reasoning.

ENDNOTES

1 Ramsey, F. P., "Last Papers" in F. P. Ramsey, *The Foundations of Mathematics*, edited by R. B. Braithwaite in the International Library of Psychology, Philosophy and Scientific Method (London 1931) Quotations according to the fourth impression 1965. For reasons unknown to me, the paper discussed here has been omitted in the revised edition, published 1978 under the title "Foundations".

2 Ramsey (1965), S. 268.

3 Ramsey (1965), S. 268.

4 Ramsey (1965), S. 263.

5 "Logic issues in tautologies, mathematics in identities, philosophy in definitions, all trivial but all part of the vital work of clarifying and organizing our thought." Ramsey, S. 264.

6 A "genuine language" is a language the terms and sentences of which have to be understood without the help of another language, especially a meta-language, which is already understood; for example, the first language a child learns, is a genuine language; but there are other cases at hand, like the introduction of a completely new scientific theory e.g. quantum mechanics, the language of which has to be grasped without the help of a previously understood meta-language.

7 A definition is, in my view, what Austin calls a "speech-act": pointing out to the reader in which way the defined term is to be understood in the future. The futuristic aspect of definitions is pointed out by Ramsey too without the "speech-act" terminology; see Ramsey (1965), S. 263.

8 Ramsey (1965), S. 267ff.

9 In his famous paper "The Relation of Sense-Data to Physics" Russell tried to do precisely this: explicating "physical objects in terms of sense-data"; see his *Mysticism and Logic* (London [10]1951), S. 147.

10 I guess that Ramsey was inspired by the writings of H. Weyl, whom he criticized in *The Foundations of Mathematics* (1925) and *Mathematical Logic* (1926) defending Russell's logicism. Later, however, he became convinced by H. Weyl that a general proposition is not a judgement, but "a sort of cheque which can be cashed for a real judgement when an instance of it occurs"; Ramsey, S. 70.

11 Ramsey, S. 252; the last three quotations are taken from part B of the "Last Papers", entitled "General Propositions and Causality".

12 There are several theories of mind, criticized by Ramsey, that fall short in this respect, e.g. Carnap's theory of mind in *Der logische Aufbau der Welt* as well as all materialistic theories of mind. The mistake is always the same: They start from facts, we know about our own mind, making "theories" about other minds in terms of protons, neurons etc. and then turn back on our own mind and claim "that what are *really* happening there are simply these theoretical processes" forgetting the hypothetical character of the theory; see Ramsey, S. 266.

13 Ramsey (1965), S. 268.

* * *

LEIBNIZ' REPRÄSENTATIONSFORM EINES ADÄQUATEN BEGRIFFES VON ‚GEIST' IM VERHÄLTNIS ZU DESCARTES UND KANT.

Michael Benedikt
Universität Wien

‚Repräsentation' heißt seit Leibniz immer ein Zweifaches: einmal die entweder kohärente oder sogar korrespondierende Vorstellung eines mannigfaltig Gegebenen in einem Modell oder Konstrukt – seither auch Darstellung oder Abbildung genannt –, anderseits jedoch die Rolle, welche ein Vorstellendes – für Leibniz auch andere Lebewesen – in dieser Ausübung bzw. auch in anderen trivialen Tätigkeiten in einem öffentlichen Verband im Verfolgen von darstellbaren Regeln zu spielen vermag. Die *Vereinbarung* dieser beiden Bedeutungen verlangt aber einen adäquaten, nicht bloß klaren oder distinkten logischen Diskurs, in welchem die Bedeutung von ‚Geist' einen gewissen, also zumindest klaren und distinkten Stellenwert hat. Meine Untersuchung ist zunächst ideengeschichtlich, hat aber auch mit einigen gegenwärtigen Problemen zu tun.

I. Gilbert Ryles Mißverständnis gegenüber Descartes' ‚Mythos'.

Zusammen mit dem späten Wittgenstein und dem frühen Heidegger distanziert Gilbert Ryle in *The Concept of Mind* (dt. *Der Begriff des Geistes,* Stuttgart 1969) eher die rationalistisch *enge* Konzeption als die beiden *weiteren* Auffassungen des Begriffes von Geist bei Descartes: Ich meine die auf den Geistbegriff des Subjektes hingeordnete Ontologie der beiden Substanzen, die des sum cogitans und der res extensa. ‚Geist' oder ‚Seele', in der Ambivalenz von ‚mens' als Modalbegriff des ‚sum cogitans', ist für Heidegger ein *wohlfundierter* Modus des zunächst pragmatisch orientierten und ausgelegten ‚In-der-Welt-Seins', für den späten Wittgenstein hingegen eine paralogistische und paradoxe Reflexionsform bzw. eine ausweglose, therapeutisch unfruchtbare Beendigung eines philosophischen Diskurses.

Die in der Diskurstheorie, besser Sprachspielkonzeption, des späten Wittgenstein distanzierte Form des cogito ergo sum erinnert aber an Kants Auffassung des beschränkten, also noch nicht ästhetisch gefaßten ‚Geistes', somit der Repräsentationsform der Verdoppelung von theoretischer Referenz, also im Sinne der Thematisierung dessen, was wir tun, wenn wir die Welt unserer Erfahrung als Objekt der Wissenschaft oder bloß als etwa in Rechtstiteln beschreibbaren Strukturkomplex vornehmen einerseits und der Strategie der Begründung dieser Referenz anderseits.

Die früheste cartesische Aufgabe, wie wir in einer einzigen analytischen Methode überhaupt erst zu dieser Referenz gelangen, was dann auch sowohl Referenz als auch Begründung bzw. Legitimation umfassen soll, ist für Descartes immer ein normatives Postulat gewesen, aber nie zur Durchführung gelangt. Für unseren Zweck der Rezeption des ‚Geistes' in der Repräsentationsform von Leibniz mit Bezug auf Descartes unterscheiden wir aber drei Etappen in der Aneignung des Begriffes von Geist, wovon Ryle sich eben nur der letzten bedient.

1. Eine frühe Konzeption von den *Regulae ad directionem ingenii* bis zur Legitimation der wissenschaftlichen Methode im *Discours de la Méthode* ist ausgezeichnet durch eine nicht eingeholte Differenz zwischen Selbstreferenz und Gegenstandsbezug (res ipsa). Diese *Differenz* heißt ‚Geist' zunächst als ‚intellectus ipse' der Vereinbarung von ‚sagacitas' und ‚perspicuitas', sodann als regelgeleitete ‚bona mens' oder ‚ingenium' mit Bezug auf die konstruierte

oder vorausgesetzte ‚res ipsa', deren Programm in der achten bis zwölften Regula entworfen wird, allerdings ohne Auflösung jener Differenz.

2. Die im *Discours* vorbereitete, in den *Meditationen* zum Durchbruch gekommene Fixierung des Unterschiedes von Reflexion und Referenz unter der Bedingung einer Einordnung deskriptiver Wissenschaft und legitimierender Theologie in Anbetracht der neuen Sprachanalyse und Urteilslehre ist der prekäre Ausweg aus dem Unvollendeten, aus den *Regulae*, im Zeichen eines erweiterten Begriffes von ‚Geist' im Sinne von ‚spiritus' (als malignus oder als summum bonum).

3. Die petrifizierte Lehre der drei Substanzen, wie sie teils in den *Meditationen* und deren Apologie, vor allem aber in den *Prinzipien* bzw. der Repräsentation im Schreiben an Picot vorgetragen ist, stellt das Gerüst des rationalistischen Geistes in der üblichen Fassung des ‚intellectus ipse' dar.

Gilbert Ryle hält sich hauptsächlich an den Übergang von der zweiten zur dritten Konzeption, wenn er diese auch in seiner Auffassung des ‚Mythos von Descartes' (Kybernetes-Gleichnis) rückinterpretierend schon in den Übergang der ersten in die zweite Phase fallen läßt. Hierbei bedient er sich aber immer des reduzierten Idioms im Sinne der die Differenz begründenden Selbstreflexion einer Einheit von ‚sagacitas' und ‚perspicuitas'. Dagegen hat im *Discours* der Vorrang des ‚rem cogitare' (also nicht des ‚se cogitare') einen ganz anderen methodischen Charakter und eine andere ontologische Herkunft als die reduktive Formulierung des substantial oder dynamisch gefaßten Bewußtseinshabitus, welche vor allem Gegenstand der Ryle'schen Kritik der kategorialen Verwechslung darstellt. Während die Aufgabe der referentiell-objektiven Synthesis die frühen Ansätze in den *Regulae* überfordert und in der Folge das Unternehmen einer Begründung der Objektivität und Mathematisierung der Naturwissenschaften vermutlich aufgrund der impliziten Bacon-Kritik abgesetzt wird, ist die Strategie im Discours nach dreierlei Hinsichten abgesichert: Nach der ontotheologischen Lösung, derjenigen des technischen Experimentes von Modellen, was deren *Erklärungswert* und *Fruchtbarkeit* in der Voraussage betrifft; zuletzt nach der Repräsentationsfunktion sprachanalytischer Relevanz im propositionalen Urteil.

Ryle ignoriert diese Verfahrensweisen und zieht sich in einer mißverstandenen Wiederaufnahme zur Zusammenfassung des 5. Discours (Leib-Seele-Problem) auf die Kulmination von zwei Abstrakta zurück (Repräsentation in der vergeblichen Einheit des Selbstbewußtseins als Rolle instrumenteller Vernunft und als Präsentation der res ipsa), von denen Spinoza schon feststellt, daß sie nur dann einem Übergang von obskurer zu klarer, wenn auch nie adäquater Erkenntnis zugänglich seien, wenn man gerade die vermittelnde Substanz einklammert. Ryles Idee der Inexistenz des Geistes als Kybernetes eines Körperschiffes, ein altes Modell, das im 5. Discours wieder aufgenommen ist, unterschlägt aber zweierlei: Die technische Rekonstruktion unserer wissenschaftlichen Objektwelt findet für Descartes an der Menschenwelt ihre akkurate Grenze, wenn diese auch im Zuge der Verwissenschaftlichung modifiziert wird. Anderseits ist die Bewußtseinseinheit unserer Referenz auf die Körperwelt bzw. derjenigen auf uns selbst als auf unseren Leib auch keine Angelegenheit der Kategorialform des subjektiven Geistes als Kybernetes, sondern Sache des ‚geistigen vinculum' des ‚deus sive natura', eben auch schon bei Descartes. Deren bloß subjektive Reflexionseinbindung ohne Kenntnisnahme des Antecedens als ‚rem cogitare' und der Konsequenz als ‚se cogitare' sind paralogistisch für den gesunden Verstand, narzißtisch für den gesellschaftlich oder individuell verfaßten kranken Verstand, fatal für die Kritik der historischen Vernunft, im ganzen höchstens ein Anstoß zur Produktion für den Geist des künstlerischen Ingeniums.

Der andere wunde Punkt Ryles ist die Verkennung von Descartes' *pragmatischem Hintergrund* der Verwissenschaftlichung einer vorläufigen Moral: Die Legitimationsstrategie der Objektion unseres wissenschaftlichen Verfahrens ist nämlich eingebettet in die Verfahrensweise des pragmatischen Geistes und dessen technische Überlebensstrategie gemäß diskursiver Logizität. Erst ein zu Ende geführtes Experiment der Verwissenschaftlichung könnte mit Bezug auf die sukzessive Ablösung unserer pragmatischen Interessen zugunsten einer Gesellschaftsethik zu so etwas wie einem Erweis des ‚category mistake' führen, wenn wir zuvor die

philosophische Methode außer acht lassen. Leider hatte sich Descartes den Primat des Praktischen durch den auch für modale und relationale Begriffsverbindungen gebrauchten Regelbegriff des Kontingenten als bloß ‚moralement possible' weitgehend verbaut.

Kant hat im Gegensatz zu Ryle versucht, dem Rechnung zu tragen, indem er anstelle des Gottesbegriffes, der beim mittleren Descartes die Einheit von Körper und Seele, die Referenz der Vorstellungen auf ihre Objekte, sogar die Einheit der Affinität von methodischen Teilschritten innerhalb von mathematischen Modellen der Wahrheitswertigkeit ein anderes Maß der wahren Propositionen zur Legitimation der Affinität vorschreibt, nämlich seine Logik der Synthesis apriori. Synthetische Sätze a priori ersetzen dann also, was bei Descartes den eigentlichen Vermittlungsbegriff des Geistes ausmacht, demgemäß die paralogistische Referenz der Seele auf sich selbst immer nur ein narzißtisches Beiwerk oder ein Umweg gewesen war, in den die faule Vernunft immer dann hineintappt, wenn sie die Bedingungen der Erkenntnis eines Sachverhaltes nicht mit Bezug auf ein primäres praktisches Erkenntnisinteresse anzugeben vermag.

Wie sich also die regelgeleitete Verbindung der Terme eines Urteils (im Sinne des Primats von Bejahung, Koordination und Optimierung) an der realisierbaren Einheit der Notionen aufweisen läßt, ohne wie Descartes, Spinoza oder Leibniz auf ein summum bonum zurückgreifen, vielmehr dies durch das regelgeleitete Handlungsinteresse zu supplieren, ist der Hintergrund von Ryles vordergründigem ‚Mythos des Geistes'.

II. Leibniz' wiederholte Aneignung der res cogitans

Der bloß beiläufige, kontingente Übergang von ‚perspicuitas' zur ‚bona mens' des ‚ingenium', die neugierige ‚curiositas', hat Leibniz zunächst in Form der Gier des Aufsuchens von Nachlaßschriften Descartes' eingeholt. Dies schlägt selbst noch in seiner Auseinandersetzung mit der Äbtissin von Herford zu Buche.

In seinem Briefwechsel bzw. den Entwürfen der Korrespondenz mit Sophie-Charlotte von Preußen hat Leibniz prägnanter zwei Hauptthesen und ein verstecktes drittes Thema vorzutragen, und hier kommt er wieder auf Descartes zurück: Zunächst geht es in der Kontroverse mit Locke um die intersubjektive Erfaßbarkeit der primären Qualitäten und der Relationen der sekundären Phänomene sowohl gegenüber den ‚petites perceptions' des Unbewußten als auch gegenüber der Variabilität der Überschüsse unserer Reflexion mit Bezug auf unsere Sinneseindrücke: Hier wird dem vergleichenden und selegierenden Gemeinsinn zwischen Reproduktion und Antizipation die geistige Tätigkeit als primärer Ordnungsfaktor vorangestellt: das, was die Sinne, sie objektiv ordnend, übersteigt. Anderseits können wir die Komplexionen der Materie in der Variabilität ihrer Konstellation – unter gegebener Konstanz des Raumes und der Materie selbst, eine zweifelsohne prekäre Annahme – nur in der Differenz einer jeweils stufenweise vergeistigten Materie und selbst einer davon noch distanzierten und von der Materie losgelösten Substanz annehmen. *Diese Differenz zu setzen* ist nun Angelegenheit der res cogitans. Hier liegen Argumente vor, die über die monadische Verfaßtheit der Naturwelt bzw. der natürlichen Gesellschaften hinaus, wie sie Leibniz schon seit 1694 durch seinen Begriff der Monade vorstellig macht, eben in der Infragestellung der „natürlichen Gesellschaften" (von Ehe über Gemeinwesen, Großfamilien bis zur Kirche) mit Bezug auf nunmehr vordringliche Formen *künstlicher Gesellschaft* zur Geltung gebracht werden.

Wenn nun in seinen *Bemerkungen zu den cartesischen Prinzipien* (1692) ebenso wie in seiner Schrift *Gegen Descartes* vom Mai 1702 sowohl gegen den Atomismus, die abstrakte Geometrisierung des Raumes, die Unterschlagung der Dynamik gegenüber der Mechanik, das Überspringen des produktiven Möglichkeitsbegriffes bzw. der beiden Optimierungsgesetze Stellung bezogen wird, so greift darüber hinaus Leibniz sogar Descartes an der Wurzel seines paradoxen Zweifelverfahrens und seines Konzepts des sum cogitans an.

Gleichwohl ist in der Vereinbarung der ‚vis primitiva activa et passiva' als Grundlegung dynamischer Prozesse von Leibniz das Problem der kontingenten Propositionen eingeführt,

deren unendliche und unabsehbare Auflösung (resolutio) nicht nur den Kontingenz-Begriff des ‚moralement possible' des Descartes wieder aufleben läßt, sondern auch dessen pragmatisch-gesellschaftlichen Hintergrund wieder zur Geltung bringt.

Das ‚moralement possible' ist als die Sphäre faktischer Möglichkeit, deren Verifikation immer nur nach bloßen Wahrscheinlichkeitskriterien im Kalkül des Ermessens vor sich geht, rezipiert: Seine Bestimmbarkeit und konkrete Bestimmung des Optimierens findet deshalb zunächst vor allem in hypothetischen Propositionen oder Handlungsanweisungen, die für einen durch sie eingrenzbaren Geltungsbereich auf Abruf funktionieren, statt. Diese sind aber auch, ohne Zweifel, der kontingente Hintergrund von Descartes' provisorischer Moral, wenn auch deren Prinzip in der vierten Maxime wieder auf eine Reflexionsnorm des ‚sum cogitans, jetzt aber in verkürzter pragmatischer Form ohne das Antecedens des ‚rem cogitare' zur Geltung gebracht ist. Sowie also Leibniz die Sphäre des Kontingenten in seine Titel des abgeleiteten ‚principium grande' übernimmt, deren Exposition und Transformation ich hier voraussetzen muß, übernimmt er zugleich die pragmatische Reflexionsbasis der von Descartes mitbedachten Sphäre der Bedürfnisgesellschaft im Sinne des Reflexionsstatuts des Primats des ‚se cogitare' vor der Affinität zum rem cogitare. Wenn also noch im Mai 1702 Leibniz gegen Descartes den Vorwurf erhebt, weder Ausdehnung noch Dauer begriffen zu haben (kontinuierliche und simultan gesetzte Wiederholung), noch auch die Sphäre des Kontinuums zu begreifen, die er selbst als reversible Manifestation eines Mannigfaltigen vorstellig macht und strukturell zwischen tätiger und passiver Kraft einbindet, so versucht er auch hier noch dem Argument Spinozas gegenüber Descartes (Einklammerung der Substanz selbst) gerecht zu werden.

Umso seltsamer nimmt sich aber einige Monate später Leibnizens Wiederaufnahme des Kerngedankens des autarken Intellektes im Übergang von der referentiellen Descriptionsform zur pragmatischen Form der Selbstmanifestation aus, ebenso wie die neuerdings erhobene Idee einer Konzeption des im Raum zusammengesetzten Kontingenten, was weder phänomenal konstruierbar noch auch harmonikal in der Dominanz einer Monade über andere und deren Übereinstimmung mit je anderen dominierenden Monaden antizipierbar ist.

Es gibt zwei Gründe für Leibnizens Rückgriff auf den hybriden autarken Intellekt Descartes': Einerseits sieht er das Problem komplexer materieller Zustände, auch mit Bezug auf organisierte Automaten, über die phänomenale Konstruktion hinaus dringlich werden. Im Zusammenspiel von vis activa und passiva ist Descartes' Strukturmodell eines Gleichgewichts von Mustern nicht auszuschließen; zugleich zeigt jedoch die Abhebung vom Konzept der Monade und dem der phänomenalen Konstruktion die Verselbständigung einer Relationssphäre an, in der die verschiedenen Elemente oder Entitäten bzw. Komplexionsmuster der Erscheinungsmannigfaltigkeit eine jeweils variable Rolle spielen können. Genau dies ist aber sowohl im Sprachkonzept des 5. Discours von Descartes ebenso wie in seiner Idee der provisorischen Moral oder aber im Konzept einer künstlichen Bedürfnisgesellschaft auf dem Boden jener Moral vorgedacht. Wenn also Leibniz ab Juni 1702 den mittleren Descartes zwischen seiner Referenztheorie der *Regulae* und dem ‚Totalexperiment' rezipiert, rezipiert er über die Transformation des pragmatischen Momentes an der *Exigenz*, also der erzwungenen Affinität von Bedeutung und Erfüllung des ‚sum cogitans', auch die Rückseite einer vorläufigen Moral der Bedürfnisgesellschaft, ebenso wie die Synthesis von vis primitiva activa und passiva eine Komplexion mehrerer Verfahrensschritte der Verselbständigung von Relationsstrukturen vorstellig macht.

III. Die späte esoterische Lösung von Leibniz

Wenn auch das monadische System des Verhältnisses von realer Möglichkeit, der Kompossibilität atomarer Tatsachen bezüglich deren Optimierung auf der Grundlage einer Disjunktion phänomenaler Konstruktion und noumenaler Ontologie schon ab 1696 konzipiert ist, tritt es erst stärker in den ontologisch-naturphilosophischen bzw. gesellschaftstheoretischen Ent-

würfen im Anschluß an die Theodizee als jetzt dominante und reife exoterische Konzeption hervor.

Parallel zu diesen Untersuchungen der universellen Harmonie der Individuation und der auch gesellschaftstheoretisch bedeutsamen Kompossibilität von Individuen (im Sinne der Gruppierung natürlicher Gesellschaften) wird jedoch schon ab 1706 das Programm einer *esoterischen* Fragestellung fruchtbar. Es ist zunächst gekennzeichnet durch den Leitgedanken der ‚dominierenden Monade' in komplexen Organisationen, führt von da zu den Kompossibilitätsmustern einer nach Wahrheitskriterien strukturell gefaßten Objektivität (res substantiata) und leitet zu den Erörterungen der Verbundformen künstlicher Gesellschaften über: Wenn es kein Mittleres zwischen phänomenaler Welt und Monadenwelt gäbe, so lautet die Primärthese, so gäbe es nur Modellkonstrukte hier und Verinnerlichungsstrategien des Geistes zugleich mit einer theologisch bedingten Harmoniekonzeption da, aber keine öffentliche gemeinsame Welt. Ausdrücklich wird dies in mehrerlei Hinsichten exponiert, wenn wir einmal die theologische Perspektive außer acht lassen, wie dies auch die jüngeren Studien hierzu von Blondel, Mathieu, Serres zeigen.

1. Die kontingenten Elemente einer Strukturkomplexion erhalten ihre Prägung nur durch die logische Synthesis der Struktur selbst – Leibniz nennt sie ‚vinculum substantiale' – nicht durch die interne Verfaßtheit bzw. durch die abstrakte Interpretation mittels eines modellhaften Charakters.
2. ‚Vinculum substantiale' als Gleichgewichtsbedingung eines komplexen Systems erlaubt die Austauschbarkeit von akzidentellen Relationsmustern.
3. ‚Vinculum substantiale' ist der bloßen Kausalreihe eines Determinismus ebenso übergeordnet, wie es eine Synthesis der bloßen Implikation gemäß dem Statut notwendiger Wahrheit kontingent modifiziert.
4. Die resultierende Komplexion ist keine geschlossene, sondern eine offene Relationspyramide, deren Entitäten hierdurch nicht mehr isolierte feststehende Monaden sind, sondern als eine im Kontinuum der Vernetzung stehende Ensemble-Mannigfaltigkeit in einer Entfaltungsstrategie dargestellt werden müssen.
5. Durch die Einführung der pragmatischen Dimension des ‚sum cogitans' in die Entfaltung der Monaden ebenso wie durch die Feststellung der Abhängigkeit des ‚ergo sum' vom ‚rem cogitare' (und nicht umgekehrt wie in der bloß biologistisch vorgestellten Explikation etwa eines Genoms) wird die kontingente Relevanz des ‚moralement possible' aus der prohibitiven Verfaßtheit natürlicher Gesellschaft in eine permissive Funktionsform ‚künstlicher' Sozialisation transformiert.

In dem 10 Jahre währenden Briefwechsel mit Des Bosses hat Leibniz diese Sphäre der Synthesis der künstlichen Gesellschaft auf dem Boden der Transformation des Hintergrundes des sum cogitans, also der provisorischen Moral, immer wieder vorstellig gemacht; die Auflösung ist allerdings zunächst eher in theologischer als in soziologischer oder ökonomischer Relevanz präsentiert. Leibniz hat damit über die Transformation des Hintergrundes des rem cogitare hinaus, also durch die pragmatische Moral des ‚moralement possible', der bloß optimierbaren Beliebigkeit, einer Regel Folge zu leisten, einen Übergang von der Naturwelt in die Menschenwelt vorgestellt.

IV. Abschluß und Vorblick

Zusammenfassend möchte ich erinnern, daß Ryles sprachanalytische und behavioristische Konzeption die logische Struktur des Primats des ‚rem cogitare' vor dem ‚se cogitare' ebenso vernachlässigt, wie deren Abwandlung von drei Hauptphasen, die er ebenso wie deren Übernahme durch Leibniz unterschätzt, besonders was die Transformation des pragmatischen Hintergrundes aus der natürlichen Gesellschaft in die künstliche der Menschenwelt betrifft.

In diesem Zusammenhang ist zu erinnern, daß der eigentliche Geistbegriff gerade auch schon seit Platos *Timaios*, wenn auch nicht in der Selbstreflexion der pragmatisch oder ideolo-

gisch isolierten ‚membra disjecta', so doch in der optimierbaren Synthesis der kompossiblen Elemente liegt. Dies besagt auch das vinculum substantiale superadditum, wenn dies auch als Geistbegriff, im Gegensatz zur ontotheologisch verfaßten Natur Platos, jetzt im Sinne eines Primates pragmatischer Reflexion vorgestellt wird, zugleich aber auch als Basis einer realistischen Referenztheorie fungiert, welche als Entlastung im Übergang von natürlicher zu künstlicher Gesellschaft und deren Gerechtigkeitsstruktur fällig wird. Ohne Zweifel hat Leibniz dieser arkanen Theorie zwischen 1706 und 1716 noch keine ausdrückliche Logik verschafft, wenn auch eine Reihe von ontologischen Entwürfen hierzu vorliegen. Diese Untersuchung hat Kant in seinem System der Synthesis a priori einzuführen getrachtet, welches die Vereinbarkeit von Handlungsteleologie der Intention von Motiven mit den Resultaten von Handlungen einerseits als auch andererseits deren Horizont einer Lebenswelt mit Blick auf die Bestimmung eben von kontingenten Sachverhalten unter Bedingung eines metatheoretischen Strukturmusters im Sinne der Gültigkeit einer wissenschaftlich akzeptierbaren Objektivität plausibel macht. Auf dem Boden der von ihm seit 1772 übernommenen Synthesis, im Primat der Relationen auch der sekundären Qualitäten, sind aber die Titel der vier Grundsätze aus der Dominanz des zureichenden Grundes im Sinne der Transformation von Regeln der Subsumtion, Bejahung, Koordination und Optimierung in den Primat der Analyse der Erfahrung auf dem Boden eines Zeitschematismus (zwischen gerichteter Zeit und der Zeitform reversibler Prozesse) transformiert. Ob jedoch jenes ‚vinculum substantiale' in der kritischen Transformation auch für die Modellbildung der ‚besonderen Analogien der Erfahrung' gemäß einem Primat der Raumschematik (und der Einbindung gerichteter Zeit gemäß entropischer Dissipation) im Sinne des ‚Fortschrittes der Physik' zureichend ist, dies ist eine andere Frage, welche ich hier ebenso wenig diskutieren will wie die Relativierung instrumentell ausgestatteter pragmatischer Vernunft durch eine ethisch verbindliche Synthesis weiter Pflichten, welche auf die Folgen der Handlung je Anderer geht. Hier erst würde jener von Descartes durch Leibniz vorgeschlagene Übergang von Natur- in Menschenwelt in letzterer einheimisch.

* * *

ZUR STRUKTUR PSYCHOLOGISCHER THEORIEN

Christfried Tögel
Bulgarische Akademie der Wissenschaften, Sofia

Das von Sneed und Stegmüller ausgearbeitete strukturalistische Theorienkonzept[1] ist bisher – von wenigen Ausnahmen abgesehen[2] – vorwiegend auf physikalische Theorien angewendet worden. Bei seiner Übertragung auf psychologische Theorien ergeben sich eine Reihe von Schwierigkeiten. So läßt sich z.B. für die wenigsten der Kern der Theorie als mathematische Struktur beschreiben.

Trotzdem sollen hier einige Erkenntnisgewinn versprechende Gesichtspunkte der strukturalistischen Betrachtungsweise psychologischer Theorien diskutiert werden.

Zu diesem Zweck seien die wichtigsten Schlußfolgerungen des Konzepts von Sneed/Stegmüller kurz resümiert. Eine der zentralen Behauptungen ist, daß Theorien gegenüber Falsifikationsversuchen immun sind, und zwar aus 3 Gründen: 1. weil die empirische Widerlegung einer Hypothese nicht mehr besagt, als daß die entsprechende Spezialisierung des Theoriekerns untauglich ist. Daraus folgt jedoch nicht, daß im Prinzip keine erfolgreichen Kernspezialisierungen möglich sind; 2. weil der Anwendungsbereich einer Theorie eine offene Menge ist, d.h. er kann abgesehen von den paradigmatischen Beispielen beliebig eingeengt werden; und 3. weil der Strukturkern einer Theorie in bezug auf eben diese Theorie relativierte theoretische Größen enthält, die eine empirische Falsifikation unmöglich machen.

Das scheinbar Irrationale dieser Behauptungen verschwindet, wenn man von der Identifikation einer Theorie mit ihren empirischen Hypothesen zu deren strenger Unterscheidung übergeht. Damit hören Theorien auf, Mengen von Aussagen zu sein, und verwandeln sich in ein – um mit Thomas Kuhn zu sprechen – „begriffliches Instrument", mit dessen Hilfe sich empirische Hypothesen formulieren lassen. So wird auch intuitiv klar, daß eine Widerlegung von Hypothesen nicht auf die Theorie selbst durchschlägt.

Ein zweiter, eng mit dem Problem der Immunität von Theorien verbundener, neuer Ansatz innerhalb des strukturalistischen Konzepts bezieht sich auf theoretische Terme. Sneed hat ein Theoretizitätskriterium eingeführt, daß auf die jeweilige Theorie relativiert ist. Eine Größe B der Theorie T ist danach T-theoretisch, wenn sie nur unter der Voraussetzung bestimmt werden kann, daß bereits eine gültige Anwendung der Theorie existiert.[3] Dieses Theoretizitätskriterium beruht nicht auf der Dichotomie theoretisch-beobachtbar, sondern auf der *Unterscheidung* der Dichotomien theoretisch-nichttheoretisch und beobachtbar-nichtbeobachtbar.

Das strukturalistische Theorienkonzept liefert noch weitere überraschende Gesichtspunkte; ich will mich jedoch auf die Immunität von Theorien, bzw. die strenge Unterscheidung von Theorien und ihren empirischen Hypothesen beschränken und zu zeigen versuchen, daß vom Standpunkt dieser Auffassung einige Probleme der Psychologie bzw. Psychologiegeschichte eine ganz natürliche Lösung finden.

Ein erstes Beispiel entnehme ich der Lerntheorie. Hier liegen am ehesten Ansätze vor, die formal gesehen mit ausgereiften physikalischen Theorien vergleichbar sind. Man denke nur an die Modelle von Estes, Bush/Mosteller und Luce. Ein Hauptproblem der Lernpsychologie besteht in der Frage, inwieweit experimentell gefundene und theoretisch abgeleitete Lernverläufe deckungsgleich sind. Für das Lernmodell von Estes versuchte man die Deckungsgleichheit durch Einführung proportionsspezifischer Reizbindungen herzustellen. Diese Methode gibt jedoch keinen Aufschluß darüber, *welche* Reizanteile wirksam sind, da die Anpassung des eingeführten Proportionalitätsfaktors durch keine empirische Aussage legitimiert werden kann.[4] Dieses Problem verschwindet nun, wenn man von der traditionellen Theorienauffassung zum strukturalistischen Theorienkonzept übergeht. Dann wird der Reiz nämlich eine

theoretische Größe in bezug auf die Estes'sche Theorie und eine empirische Hypothese besteht in der Behauptung, daß die Theorie auf einen bestimmten Lernverlauf angewendet werden kann. Erweist sich eine solche Hypothese als falsch, muß entweder der Lernverlauf aus dem Anwendungsbereich ausgeschlossen werden oder die entsprechende Spezialisierung des Estes'schen Modells verworfen werden. In jedem Falle jedoch ist die Nichtdeckungsgleichheit experimentell gefundener und theoretischer abgeleiteter Lernverläufe kein Problem der Theorie selbst.

Als zweites Beispiel für eine Neuinterpretation psychologischer Theorien vom Standpunkt des strukturalistischen Theorienkonzepts möchte ich die Psychoanalyse wählen. Es ist bekannt, daß Freud 3 Bedeutungen von Psychoanalyse unterschieden hat; nämlich 1. Psychoanalyse als Verfahren zur Untersuchung vorwiegend unbewußter seelischer Vorgänge, 2. Psychoanalyse als Behandlungsmethode neurotischer Störungen und 3. Psychoanalyse als System psychologischer und psychopathologischer Theorien.[5] Uns interessiert hier selbstverständlich nur der dritte Aspekt.

Seit fast 70 Jahren werden immer wieder Experimente gemacht, deren erklärtes Ziel die empirische Bestätigung bzw. Widerlegung der psychoanalytischen Theorie ist.[6] Eine Vielzahl dieser Versuche ist negativ ausgegangen, d.h. viele psychoanalytischen Theorien werden als empirisch widerlegt angesehen. Wieso, so muß man sich fragen, hat diese Entwicklung nicht zu einem Niedergang der Psychoanalyse oder doch wenigstens zu einer Abnahme ihrer Anziehungskraft geführt. Sicher liegt ein Grund in der meisterhaften Darstellung des Gegenstandes durch Freud selbst. Allerdings ist es eher seine schriftstellerische Begabung als die innere Folgerichtigkeit der Dinge, die dem Leser eine eiserne Logik suggeriert. Doch Freuds Stil kann nicht alles erklären. Meines Erachtens liefert das strukturalistische Theorienkonzept eine Erklärung des scheinbar widersinnigen Verhaltens vieler Anhänger der Psychoanalyse. Nimmt man nämlich mit Sneed und Stegmüller an, daß nicht Theorien, sondern nur deren empirische Hypothesen getestet werden können, so liegt in dem Festhalten an psychoanalytischen Theorien trotz Widerlegung psychoanalytischer Hypothesen nichts Irrationales mehr.

Ich möchte diesen Gedanken anhand Freuds Persönlichkeitstheorie, wie sie in seiner Schrift „Das Ich und das Es" aus dem Jahre 1923[7] niedergelegt ist, erläutern. Kern dieser Theorie ist die Idee vom strukturellen Aufbau der Persönlichkeit, bestehend aus Es, Ich und Über-Ich. Als paradigmatisches Beispiel dieser seiner Schichtenlehre betrachtete Freud die Neurosenätiologie.[8] Danach erkrankt der Mensch an dem Konflikt zwischen den Ansprüchen des Trieblebens und den Widerständen und Verboten die das Ich bzw. das Über-Ich ihnen entgegensetzt.

Zuerst muß bemerkt werden, daß das Freudsche Schichtenmodell einer empirischen Überprüfung schon aus dem Grunde unzugänglich ist, weil sich keine experimentellen Bedingungen angeben lassen, die es verifizieren bzw. falsifizieren könnten. Dieser vom Standpunkt der traditionellen Theorienauffassung entscheidende Mangel wird unter dem Blickwinkel des strukturalistischen Theorienkonzepts zu einem wesentlichen Chatakteristikum der Theorie. In bezug auf das psychoanalytische Persönlichkeitskonzept erweisen sich darüber hinaus Es und Über-Ich als theoretische Begriffe und es existieren keine vom Schichtenmodell unabhängige Methoden ihrer Bestimmung. Folglich hatte es Freud gar nicht nötig, seine Theorien – etwa durch Einführung des Widerstandsbegriffs – zu immunisieren. Sie sind einfach immun.

Hat dann – so erhebt sich die Frage – empirische Überprüfung überhaupt noch einen Platz in der Wissenschaft? Die Antwort lautet: ja, und zwar einen sehr wichtigen Platz. Empirisch überprüft werden nämlich Anwendbarkeitsbehauptungen bestimmter Kernspezialisierungen, d.h. empirische Hypothesen. Im Falle der Bestätigung wird 1. der Anwendungsbereich der Theorie erweitert und 2. die begriffliche Struktur der Theorie durch Aufnahme der entsprechenden Kernspezialisierung differenziert. Umgekehrt wird bei Widerlegungen der Anwendbarkeitsbehauptung das entsprechende Gebiet aus der Menge der intendierten Anwendungen ausgeschlossen und die Kernspezialisierung verworfen. Auf den Kern der Theorie selbst schlägt die Widerlegung jedoch nicht durch.

In bezug auf das psychoanalytische Strukturmodell hat Freud selbst eine Kernspezialisie-

rung vorgeschlagen und ihre Anwendbarkeit auf schizophrene Psychosen behauptet. Danach entstehen diese aus einem Konflikt zwischen Ich und Außenwelt.[9] Da hier schon die theoretischen Begriffe „Es“ und „Über-Ich“ eliminiert sind, ist eine empirische Überprüfung im Prinzip möglich.

So zeigt sich auch auf dem Gebiet der Psychologie, daß vom Standpunkt des strukturalistischen Theorienkonzepts das Festhalten an Theorien trotz Widerlegung ihrer empirischen Hypothesen seine scheinbare Irrationalität verliert.

ANMERKUNGEN

1 Sneed, J., *The Logical Structure of Mathematical Physics* (Dordrecht 1971); Stegmüller, W., *Theorienstrukturen und Theoriendynamik* (Heidelberg-Berlin-New York 1973); Stegmüller, W., *Neue Wege der Wissenschaftsphilosophie* (Heidelberg-Berlin-New York 1980).

2 Z.B. W. Diederich und H. F. Fulda, Sneed'sche Strukturen in Marx' „Kapital“, *Neue Hefte für Philosophie* 13 (1978), S. 47–80.

3 Vgl. dazu W. Stegmüller, *Hauptströmungen der Gegenwartsphilosophie*, Bd. 2 (Stuttgart[6]), S. 474ff.

4 Vgl. zu diesem Problem F. Klix, *Information und Verhalten* (Berlin 1973), S. 459ff.

5 Freud, S., „Psychoanalyse“ und „Libidotheorie“, in: M. Marcuse (Hrsg.) *Handwörterbuch der Sexualwissenschaft* (Bonn 1923); vgl. auch J. Laplanche und J. B. Pontalis, *Das Vokabular der Psychoanalyse* (Frankfurt 1973).

6 Vgl. dazu F. Kiener, Empirische Kontrolle psychoanalytischer Thesen, in: *Handbuch der Psychologie*, Bd. 8, 2. Halbb. (Göttingen 1978).

7 Freud, S., Das Ich und das Es (Wien 1923).

8 Freud, S., *Neue Folge der Vorlesungen zur Einführung in die Psychoanalyse*, in: Studienausgabe Bd. 1 (Frankfurt 1969), S. 496.

9 Freud, S., *Der Realitätsverlust bei Neurose und Psychose*, in: Studienausgabe Bd. 3 (Frankfurt 1975), S. 355–361.

* * *

HOLISTIC BEHAVIORISM

Edward F. Becker
University of Nebraska, Lincoln

Our belief that other people are conscious beings is evidently based upon our observations of their behavior. If asked to say why we thought someone was feeling pain or was thinking of Vienna, we would cite some bit of behavior, such as a grimace, or something the person had said. But does the observation of the behavior really give us a good reason for believing in the conscious processes, which, after all, we never directly observe (except in our own case)? This is the problem of "other minds."

The purpose of this paper is to sketch a solution to this problem. Because space is limited many details will have to be passed over, but hopefully the main outlines of the proposed solution will become clear.

I

All mental states are either *occurrent* or *non-occurrent*. Occurrent mental states include sensations, thoughts, feelings, emotions, and in general any mental states which occur at, or during, a definite period of time. Non-occurrent mental states include beliefs, desires and intentions, as well as moods (such as depression) and character traits (such as pride). They come and go, but not at any definite time, and, unlike the occurrent states, need not involve any conscious processes.

Often, when we know about the occurrent mental state of another it is because he has *told* us how he feels, or what he is thinking. There is, of course, room for error in such cases: our interlocutor may be dishonest. The same applies to non-verbal evidence for mental occurrences. How are we to know that a person is not feigning anger, or pain, or sadness?

Well, we might know these things, in a reasonably strict sense of "know", if we knew that the person was honest, and thus that it would be out of character for him to deceive us by word or deed. At this point it may seem that we have merely traded the problem of recognizing mental occurrences for the problem of recognizing honest people. There is, however, a certain benefit to this exchange. Honesty is a non-occurrent mental state. (Specifically, a disposition to avoid deceptive behavior.) Thus, if we could solve the general problem of recognizing non-occurrent mental states, we would, in particular, have learned how to recognize honesty, and would thereby have resolved, at least to a large extent, the problem of inferring mental occurrences from behavior.

Knowledge of mental occurrences, then, depends on knowledge of non-occurrent mental states. It is on knowlege of the second kind, therefore, that we should focus our discussion.

II

One proposed account of our knowledge of non-occurrent mental states is analytical behaviorism. On this view, statements about such states are to be analyzed as statements about a person's behavior or his behavioral dispositions. Irrascibility, for example, might be explained as a disposition to engage in angry behavior. If analytical behaviorism were true our knowledge of non-occurrent mental states would be grounded in observations of behavior because every statement about the former would be equivalent to some statement about the latter.

Apart from whatever difficulties may attend particular behavioral analyses, there is a general problem which casts doubt on the whole behaviorist program. There are certain mental states, such as belief, which, although they do *involve* behavioral dispositions, are such that *which* dispositions they involve depends upon the person's *other* beliefs, desires, etc. For example, suppose that while playing poker I come to believe that one of my opponents, Smith, has two pair. No doubt this belief involves certain dispositions to behave, but which ones? The answer depends upon my beliefs about the holdings of the other players in the game, about Smith's susceptibility to being bluffed, about the susceptibility of the other players to being bluffed, etc.; it also depends, obviously, upon various of my desires and intentions. There is thus no saying what dispositions my belief involves without referring to some of my other beliefs, desires and intentions. This being so, no behavioral analysis can eliminate reference to these mental states.

It seems, then, that we must reject analytical behaviorism. On the other hand, it seems undeniable that we sometimes *do* know what others believe, desire, or intend and that this knowledge is based solely on our observations of their behavior. There thus arises the *paradox of behaviorism*: it seems obvious that behavior *is* the basis of our knowledge of others' mental states; and yet, given the failure of analytical behaviorism, it is hard to see how it *could* be the basis of such knowledge.

III

We can resolve this paradox by adopting a view of the mental which may be called "holistic behaviorism." On this view knowledge of a person's mental states is derived not by inferring particular beliefs, desires, etc., from particular bits of behavior, but by attributing to him the most rational *system* of beliefs, desires, etc., which is compatible with the whole pattern of his behavior as we have so far abserved it. Further, the system of beliefs, desires, etc., which a person really has, *is* just the most rational system compatible with the totality of his behavior. There can be no basis for ascribing different mental states to people who behave in the same way. The mind, as a whole, supervenes on behavior.

Let us set this view out in more detail. Holistic behaviorism involves the following theses:

1) Statements about particular mental states cannot be analyzed in terms of behavior.

2) Nevertheless, statements about mental states *are* confirmed on the basis of observations of behavior.

3) In ascribing mental states to people we try to ascribe to them the most rational system of beliefs, desires, etc., and of degrees of belief and desire, which is compatible with their behavior.

4) Claims about individual mental states are justified only insofar as they are part of our best systematic account of a person's mental states; such claims can only be established or refuted in the context of a theory.

5) We ascribe non-occurrent mental states, though not mental occurrences, to ourselves on the same basis on which we ascribe them to others.

6) In some cases people may believe, desire, intend, etc., things which they have never thought of or which they would even deny believing, desiring, or intending.

7) The ascription of any particular mental state to a person can be made consistent with his behavior insofar as we are willing to adjust our ascriptions of *other* mental states to him. (Think of paranoia.)

8) A person's behavior is the *only* basis for ascribing non-occurrent mental states to him.

9) In ascribing systems of mental states to a person we are guided by the principle of maximum charity; we assign the most rational system of mental states compatible with the agent's behavior.

10) The agent's rationality consists in his avoiding inconsistency and acting in such a way as to maximize the expected value of his acts.

Holistic behaviorism has many things to recommend it. Here are some of them:
A) It explains the failure of analytical behaviorism.
B) It is compatible with the evident fact that observations of behavior are the basis for the attribution of mental states.
C) Unlike some accounts of our knowledge of the mental it allows for the possibility of our being wrong, in any particular case, about a person's mental state.
D) It implies the intuitively acceptable conclusion that we often do have probable (sometimes highly probable) knowledge of other people's mental states.
E) It explains, by way of the principle of charity, why everyone who understands them believes certain obvious principles of logic and mathematics. (E.g., that 1,000,017 + 1 = 1,000,018.)
F) It explains the possibility of "unconscious" beliefs, desires, and intentions.
G) It explains how, as in cases of paranoia, bizarre views about people's mental states may be irrefutable, in the sense that they do not contradict any of the behavioral data.
H) It allows reasonable scope for indeterminacy in the attribution of mental states. A person's mental states will be indeterminate when the application of the principle of charity to the observed behavior yields no uniquely best theory.

IV

In recent years, philosophers, largely because of the influence of W. V. Quine, have come to realize that establishing a scientific theory does not consist in the confirmation of isolated hypotheses by isolated bits of data. It is rather our entire theory of the world, or at least a substantial chunk of it, that has to be reconciled with our experience. Yet there has been a curious failure to extend this insight to the philosophy of mind. Even in Quine's writings there is a yearning, predictably unfulfilled, for reductive analyses in the spirit of analytical behaviorism. If what has been said in this paper is correct, such analyses are both impossible and unnecessary. Behavior is the basis of our knowledge of the mental, but not because isolated mental states can be equated with isolated bits of behavior. In understanding a person's mind, we treat his beliefs, desires, and intentions as a corporate body.*

ENDNOTE

* The influence of Quine's work on the ideas presented here will be obvious. Donald Davidson has also defended a version of holism which resembles, in some respects, the position taken in this paper. See his "Belief and the Basis of Meaning, *Synthese*, 27 (1974), pp. 309–323, and reprinted in his *Inquiries Into Truth and Meaning*, (Oxford and New York, 1984) pp. 141–54.

* * *

ÜBERLEGUNGEN ZUR ENTSTEHUNG DER LEIB-SEELE-UNTERSCHEIDUNG UND IHREM ZUSAMMENHANG MIT FRAGEN DER ERKENNTNISTHEORIE

Josef Marschner
Universität Linz

An die Unterscheidung von Leib und Seele wurde in der Geschichte der Philosophie eine Reihe von Fragestellungen geknüpft. Hier aber soll die zunächst unauffällige Tatsache beleuchtet werden, daß eine solche Unterscheidung überhaupt stattgefunden hat. Es ist wissenssoziologisch zu klären, warum die Leib-Seele-Unterscheidung (im weiteren: LSU) besteht. Mit K. R. Popper kann man davon ausgehen, daß die LSU der Versuch einer Erklärung der Erfahrung unseres Bewußtseins ist.[1] Dazu muß aber gezeigt werden, warum die Annahme der Existenz einer Seele überhaupt eine Erklärung (für die Wahrnehmung) des eigenen Subjektseins bieten kann. Eine Theorie, die dies leistet, stammt von G. Dux, und soll als Grundlage für die Beantwortung unserer Frage herangezogen werden.[2]

Dux untersucht den Prozeß des Erwerbs der Schemata, unter denen der Mensch Wissen überhaupt gewinnen kann. Seine Voraussetzung ist dabei, daß diese aus der Erfahrung stammen und in einem umfassenden Lernprozeß aufgebaut werden müssen. Mit ihnen entsteht gleichzeitig Subjektivität als Reflexivität, da der Organismus, geleitet vom Interesse der Bewältigung der Umwelt, seine Motorik steuern lernen muß.[3] Entscheidend ist nun für die Gestalt der bewußten Formen der Erfahrung, daß die Umwelt für das Kind zentral durch Subjektivität repräsentiert wird. Dux schreibt:

> Eben weil das primäre und schlechterdings dominante Objekt, mit dem das Kind umzugehen lernt, die sogenannte Bezugsperson ist, (...), ist das primitive Objektschema das des Subjekts, das Ereignisschema das der Handlung. Das Kind bildet mit anderen Worten ein Objektschema aus, in dem die Objekte ein Aktionszentrum haben wie Subjekte. Für es sind alle Objekte so gebaut, als könnten sie handeln, dies oder jenes tun und lassen. Die Konsequenzen für den Aufbau der Außenwelt sind gleichermaßen weitreichend für das Ereignisschema: Das Ereignisschema wird als Handlungsschema aufgebaut. Fortan werden Ereignisse so wahrgenommen, als hätten sie in der Binnenlage eines Objekts ihren Anfang und wären ein vom Willen dirigiertes, durch den Willen beeinflußbares Tun.[4]

Die Welt wird also in einem subjektivischen Schema begriffen. Es erklärt Vorgänge dadurch, daß diese auf einen Ursprung zurückgeführt werden. Dieser wird subjektiviert und substanzialisiert.[5] Setzt man Ontogenese und Phylogenese gleich, dann bietet sich für das Zustandekommen der LSU die Erklärung: Der Mensch muß seine eigene Subjektivität, wenn er sie sich verständlich machen will, so deuten, daß die Bedingung seines Handelns, Planens, Denkens und Träumens, seines gesamten Lebensausdrucks, selbst noch in einem gleichsam innenliegenden Ursprung, in einem substanzialisiert gedachten Zentrum, eben der Seele liegt.

Ein Aspekt ist meines Erachtens noch nachzutragen: Wenn wir annehmen, daß der Mensch sich nur auf die angeführte Weise deuten konnte, dann dürfen wir nicht unsere neuzeitlichen Vorstellungen von dem, was es heißt, Subjekt zu sein, dem archaischen Menschen zuschreiben. Denn während wir uns als die Gestalter unserer Welt erleben, stellt sich der archaische Mensch so dar bzw. handelt so, als würden die Götter vermittels der Seele ihm alle Kraft und Handlungsmacht verleihen.[6] Warum orientiert sich nun das archaische Denken in letzter Instanz am Handeln von Gottheiten? Weil diese durch die Logik dieser Weltauffassung erzwungen werden. Götter werden geschaffen durch die interpretative Umsetzung des subjektivischen Schemas in der Wirklichkeit.[7] Sie sind Ursprung und somit auch letzte Erklärung für das, was in der Seele vorgeht; am Beginn der abendländischen Kultur mindestens noch zu

jener Zeit, in der die Seele noch nicht als von sich aus formende Kraft und selbsttätiges Prinzip aufgefaßt wurde, wie etwa bei Aristoteles. Solche Ideen standen jedenfalls nicht am Anfang. Dies führt zu dem Schluß, daß die Erklärung menschlicher Subjektivität eine Geschichte aufweist. Sie hat einen Richtungssinn: Er verläuft so, daß der Mensch seine Handlungsmächtigkeit und Erkenntnisfähigkeit immer mehr sich selber zuzuschreiben beginnt.[8]

Im folgenden sollen einige Erläuterungen zum Zusammenhang der Auffassung von der Seele oder dem Geist mit der Erkenntnistheorie gegeben werden. Die Philosophie ist in der Erkenntnisfrage von einer Ausgangslage geprägt, die bei Parmenides gebildet wird. Seine Ontologie erzwingt eine Auffassung von der Erkenntnis, die die Welt, wie sie uns in der äußeren Erfahrung gegeben ist, nämlich als ständige Veränderung, als bloßen Schein durchschlägt. Nur das reine Denken vermag das wahre Sein, welches ungeworden und unveränderlich in sich selbst verharrt, zu erfassen. Diese Trennung zwischen der Sinneserfahrung und der Vernunfterkenntnis führt zu Platons Ideenlehre, derzufolge Erkenntnis darin besteht, daß sich die Seele an die Ideen erinnert, die sie geschaut hat, bevor sie in den Körper eingeschlossen wurde (Anamnesis). Die Ideenwelt als die eigentliche Wirklichkeit, als das wahrhaft Seiende gegenüber der empirischen Welt, und der menschliche Geist oder die menschliche Seele als Zugang zu dieser Welt: Dies ist der einflußreichste Erkenntnismythos bis zur Entstehung des neuzeitlichen Empirismus. Auch bei Aristoteles, den man häufig Platon entgegensetzt, geschieht Erkenntnis durch das Erfassen der Form (morphè) des Seienden, also von etwas Nichtstofflichem durch den denkenden, vernünftigen Teil der Seele. Solange die Anahme sich durchhält, es existiere eine substanzhafte Seele oder ein solcher Geist, wird dieser(m) gleichzeitig das Erkenntnisvermögen zugerechnet. Damit verbunden ist das Konzept, daß letzte, essentialistische Erkenntnis nur durch den Einsatz dieses Vermögens zu gewinnen sei. Erkenntnis wurde als ein nichtmaterieller Akt aufgefaßt, dessen ihn ermöglichender Ursprung in einer Substanz von gleicher Beschaffenheit zu suchen sei. Daß Erkenntnis dadurch verstehbar wird, daß sie auf das Fundament, dem sie entspringt, zurückgeführt wird, setzt voraus, daß Erklären überhaupt dem subjektivischen Schema als einen Ursprungsschema folgt. Im neuzeitlichen Denken wird diese eigentümliche Logik erst dort schließlich zerbrochen, wo die Philosophie dem Modell der neuen Naturwissenschaften zu folgen beginnt, welches einem nichtsubjektivischen Schema der Wirklichkeitsdeutung verpflichtet ist. Mit dem Zurückdrängen der Substanzideen etwa bei Locke und Hume erfolgt der Aufbau von Erkenntnistheorien, die gleichsam die Mechanik des Zustandekommens interner Repräsentationen erklären wollen.

Auch bei Kant findet sich die Zurückweisung einer substantialisierten Auffassung vom erkennenden Ich. In dem „transzendentale Dialektik“ benannten Abschnitt der *Kritik der reinen Vernunft* behandelt er unter anderem die Vorstellungen der sogenannten rationalen Psychologie von der Substantialität der Seele bzw. des denkenden Wesens. Ein ‚reines Ich‘ ist für ihn nicht mehr möglicher Gegenstand der Erfahrung. Es müßte es aber sein, wenn die Kategorie der Substanz darauf angewendet werden soll. Vielmehr ist es transzendentale Bedingung von Erkenntnis.

Das Vorgehen Kants, dem theoretisch Denkbaren eine Grenze zu ziehen, ist mit dem Projekt des *Tractatus logico-philosophicus* von Wittgenstein verglichen worden. Allerdings ist der Unterschied im Verständnis des ‚reinen Ich‘ gravierend: Denn während Kant der transzendentalen Subjektivität gewisse Vernunftleistungen einschreibt, und diese auf dem Weg einer transzendentalen Selbstreflexion sichtbar machen will, behandelt Wittgenstein das metaphysische Subjekt in einem viel radikaleren Sinn als Grenze. Er kann über es nichts sagen. Er muß es auch nicht, weil bei ihm die Weltkonstitution nicht aus einem wie immer gearteten Subjekt erfolgt, sondern aus der logischen Struktur der Sprache selbst. Dadurch wird Wittgensteins Ansatz im Traktat zur Grundlage der Wissenschaftsphilosophie des „Wiener Kreises“, welche ihm darin folgte. Erst nach der Aufhebung der Reduktion der Sprachanalyse auf rein syntaktische und semantische Formen und mit der Hinwendung zur Pragmatik erscheint in der neueren Wissenschaftsphilosophie bei der Erkenntnisanalyse wieder das ‚Ich‘, in Gestalt des empirischen Sprachbenützers.

Ein Punkt ist noch zu vermerken: Sowohl bei der LSU wie auch beim Erkenntnisproblem überschneiden sich eine subjektivistische und eine objektivistische Perspektive; T. Nagel hat auf sie hingewiesen.[9] Dies bedeutet, daß man, wenn man den subjektiven Zugang wählt, auf den Weg einer Phänomenologie der ‚inneren' Erfahrung gerät. Diesen Weg ist paradigmatisch Descartes gegangen. Er erzwingt metaphysische Voraussetzungen, und es ist nicht zu sehen, wie man ihnen ausweichen könnte. Im Fall eines objektivistischen Lösungswegs muß unterschieden werden: Im Leib-Seele-Problem stößt man auf Fragen von der Art, wie etwa Beschreibungen subjektiver Erlebnisse übersetzbar sind in Sprachen, die eine intersubjektive Prüfbarkeit gewähren. Das extremste Beispiel dafür ist der Behaviourismus. Was aus der subjektivistischen Perspektive gegen einen solchen Zugang spricht, ist die Art, wie wir uns selbst erleben. Dies allerdings kann kein letztes Kriterium sein. Im Falle eines objektiven Zugangs zum Erkenntnisproblem gibt es seinerseits zwei Dimensionen: Bei der ersten stellt man die Frage nach der Rechtfertigung von Erkenntnisansprüchen als Frage nach der Eigenschaft jenes Mediums, in dem Erkenntnis sich objektiviert, also in der Sprache und im handelnden Umgang mit dieser. Dies kann von der analytischen Philosophie geleistet werden. Die zweite Dimension umfaßt eine naturwissenschaftliche Untersuchung des organischen Substrats des Mediums von Erkenntnis. Dieser Weg wird von der Gehirnforschung, Neuropsychologie und Biologie beschritten. Bei ihm jedoch muß von der Rechtfertigung von Erkenntnis abgesehen werden.

ANMERKUNGEN

1 Popper, K. R. und Eccles, J. C., *Das Ich und sein Gehirn* (München ³1984), S. 199.
2 Dux, G., *Die Logik der Weltbilder* (Frankfurt a. M. 1982).
3 Dux (1982), S. 85f.
4 Dux (1982), S. 94f.
5 Dux (1982), S. 123.
6 Dazu B. Snell, *Die Entdeckung des Geistes* (Göttingen ⁵1980), S. 28.
7 Dux (1982), S. 187.
8 Snell (1980), S. 29.
9 Nagel, T., *Über das Leben, die Seele und den Tod* (Königstein 1984), S. 215ff.

* * *

WHAT MENTAL STATES ARE REALLY LIKE[1]

Natika Newton
Long Island University, Greenvale, New York

Eliminative materialism claims that folk psychology is a false theory, whose elements are not reducible to neurology because they do not pick out natural kinds.[2] If this claim is to be plausible, we must be able to see how we could be wrong in thinking the most resistant mental properties, intentionality and consciousness, to be irreducible or ineliminable. In this paper I suggest how we might be wrong. I propose that all mental concepts are metaphors for the functions of our bodies, and that all mental events actually consist of memory traces of these functions. The metaphors lead to philosophical perplexities when they are taken literally. If the proposal is right, the consequence is a blend of reductionism and eliminativism.

Reductionism and eliminativism are versions of the materialistic theory which holds that the concepts and laws of physical science apply to all aspects of human beings (contrary to Cartesian dualism). Both versions presuppose scientific realism—that is, the view that the physical world exists independently of the observer and that our best approach to this world is through the methods of natural science. Reductionism claims that concepts in folk psychology refer to natural kinds and are reducible to concepts of neuroscience (whether or not they ever are or should be so reduced), whereas eliminativism holds that these concepts will be eliminated when psychology reaches scientific maturity. My discussion of these two forms of materialism presupposes nothing about the ultimate metaphysical nature of reality, or about how or whether this could be known.

To save time I will start with an assumption: that the reducibility of neurology to chemistry and physics is not so problematic as the reducibility of psychology to neurology. Most current nonreductionist psychologies which I oppose argue almost exclusively against the latter.[3] I make this assumption because I will argue that actual mental processes, not necessarily those of folk psychology, can be reduced to biological processes which do not confront us with a mind-body problem.

1. THINKING IS LIKE ACTING ON SEEN OBJECTS. Our ordinary understanding of the mind is of a being which literally behaves just like the physical body. First, there are mental and physical 'acts.' Most mental acts, like most physical ones, essentially involve objects. Physical acts of approaching or avoiding objects, manipulating them, making and altering them, or just focussing one's eyes on them, have counterparts in the mental acts of concentrating on or repressing thoughts of objects, analyzing them, creating or developing them, or just imagining them. Mental acts can be occurrent or dispositional, like physical acts. The mind, like the body, can take an attitude or stance with respect to something. Second, there are objects in both cases. Mental objects, like physical ones, are the goal or focus of some mental action or awareness. Third, there is mental and physical perception. The most common mental metaphors describe consciousness and knowledge as seeing. Like our eyes, our minds can 'focus' on only one thing at a time. Also common are mental 'feelings' or 'sensings.' Finally, mental and physical objects are given values, which motivate action. Propositions are true or false; arguments valid or invalid; decisions good or evil. These values are offered as justification for the respective mental acts (or propositional attitudes) of affirmation or denial, acceptance or rejection, approval or disapproval.

Thus, as traditionally described, our minds parallel our bodies. They show the same abstract action patterns and the same types of relations to objects. They have goals, suffer pain (sorrow) and pleasure (joy), experience chronic conditions (moods), have special abili-

ties (intelligence or talent), and can be tired (depressed) or energetic (manic).[4] Perhaps all mental properties have physical parallels, and are nothing but neurological traces of physical action, described metaphorically.

An important objection is that in my comparison of mental and physical action, perception, etc., I have not distinguished between the mental and the nonmental, since the physical actions described are all intelligent, goal-directed actions. But as I indicated earlier, explaining goal-directedness *per se* is not the real problem in philosophy of mind. As Dennett implies, there is a continuum in biology from simple to complex, well-designed goal-directed systems.[5] The problem for such systems is that of how biology is reducible to chemistry and physics, not the problem of mind and body. The essential differences between the mental and physical acts described are that the physical acts are (a) all directed at existent objects, and (b) do not presuppose self-awareness but only object discrimination. The mental acts (a) do not require that their objects exist, and (b) sometimes presuppose self-awareness. Only in the mental acts are found the properties that have traditionally resisted materialistic explication: intentionality (in Brentano's sense) and self-consciousness or conscious awareness of qualitative states. Thus I have to show that both intentionality and self-consciousness have explicit parallels in processes or states of our physical bodies, from which their properties are entirely derived.

2. INTENTIONALITY. According to Brentano[6] mental processes are characterized by "direction upon an object", even though the object need not exist. If a thinker thinks about something only the thinker need exist. Relations between physical things, or "strict" relations, require that both relata exist. Mental relations are not reducible to physical ones.

Let us consider an example. Suppose that yesterday I was bitten by a dog and today I am remembering or imagining the event by reactivating memory traces laid down yesterday. Yesterday I had a real physical relation to a real dog, and today I have an imaginary physical relation (being bitten) to an imagined, possibly now nonexistent, dog. Both the object and the relation are imagined. If imagining or remembering can be simply a matter of reactivating memory traces in the brain, then there is at least one way I can think about something that can be analyzed as a physical type of event.

It can be objected that if imagining is just reactivating memory traces it is not propositional and hence has no semantics. Only if I imagine *that* I was bitten by a dog can the process be analyzed propositionally. But if imagining, in my example, is imagining *that*, then it is intentional in a purely mental sense since it is more than just reactivating memory traces. On this view only if I analyze imagining as mechanical will my example work, but then I have not analyzed an intentional process.

This objection brings out a key point in my analysis: when I think about a dog by imagining being bitten by it, the object of my thought is dependent on the mode of representation in the same way it would be dependent on a description. The reactivated memory traces include not only the perceptual data that was present at the time, but also my entire set of responses: muscular, proprioceptive, etc. These responses pick out one aspect of the event from other possibilities: I responded to the dog's teeth, not to his tail. The reactivation of the traces of all these responses determines what it is I imagine. My claim is that these responses are no different in kind from other biological goal-oriented processes, such as a neural crest cell migrating to its destination in the developing embryo, or aplysia recoiling from a probe.[7]

3. SELF-CONSCIOUSNESS. As intentionality can be analyzed primarily in terms of physical action, self-consciousness can be analyzed in terms of vision. A pain as the object of my attention is, I claim, an illusion concocted from two concurrent states appearing to form a whole 'object.' First there is the physical sensation. Second there is my imagining that I am looking at an object by reactivating schematic imagery of focussing my eyes, where the sensation seems to be incorporated into the imagined seeing as the object focussed on. What makes the process seem so mysterious is that I have no idea of the occurrence of the eye focus ima-

gery (even though the metaphorical language I use to describe my experience might suggest it).

In the case of more generalized self-awareness the structure of the process is the same. I am now conscious that I am sitting at my desk. First there are all my physical sensations, and in addition visual data. Second there is the act of imagining seeing a person in this place. The imagery is not necessarily of a person seen from a distance; its primary material is my current sensations which appear to form an observed whole when I activate schematic memory traces of focussing my eyes. Feedback from this process may be incorporated into the collection of sensations of which I am aware, so that my awareness gives the illusion of reflexivity.

If self-consciousness is just imaged eye focus occurring along with our sensations, how can we pay selective attention to some, ignoring others? A representation or image of eye focus can take a second, concurrent brain event as its 'object' by altering and adjusting in response to this brain event and not others. When I am 'focussed on' or 'paying attention to' certain inner states, my imagery of observation responds to changes in these states in the same way that my eyes respond selectively to patterns when watching a friend walk away into a crowd. This sort of selective responsiveness is found in biology at the cellular level.

That consciousness involves something more than eye focus imagery plus selective responsiveness to other physical states is an understandable illusion, arising because what responds to certain brain states is *eye focus* imagery. Thus it seems exactly as if I am following something with my eyes, but I am, as it were, only going through the motions.

There are many further issues that should be discussed, for which there is not space in this paper. They include the question of what causes me to notice a certain inner state, apart from the mechanism by which I accomplish this; the question of consciousness in the congenitally blind; the question of the evolutionary and cultural origins of intentionality and consciousness;[8] and the current evidence of action schemata in the brain.[9] These issues are treated elsewhere.[10]

ENDNOTES

[1] I wish to thank Eric Walther for invaluable advice and encouragement.

[2] Churchland, Paul, "Eliminative Materialism and the Propositional Attitudes," *Journal of Philosophy* LXXVIII, 2 (1981).

[3] See, e.g., Jerry Fodor, *The Language of Thought* (Cambridge, Mass. 1979), Introduction; and Hilary Putnam, "Philosophy and Our Mental Life", in N. Block, *Readings in Philosophy and Psychology* (Cambridge, Mass. 1980). Vol. I, pp. 134–143.

[4] Julian Jaynes has given a thorough and persuasive analysis of the metaphorical nature of the language of mind, and of how mental events are analogues of physical events, in *The Origin of Consciousness in the Breakdown of the Bicameral Mind*, (Boston 1976), Ch. 2. See also Margaret Boden, "Implications of Language Studies for Human Nature", in Thomas W. Simon and Robert J. Scholes, eds., *Language, Mind and Brain* (Hillsdale, N. J. 1982).

[5] Dennett, Daniel, "Intentional Systems", in *Brainstorms* (Cambridge, Mass. 1978), pp. 8–9.

[6] Brentano, Franz, *Psychologie Vom Empirischen Standpunkte* (Leipzig 1924), I, pp. 124–125. Selections translated in R. M. Chisholm, ed., *Realism and the Background of Phenomenologie* (Glencoe, III. 1960).

[7] Jacobson, Marcus, *Developmental Neurobiology* (New York 1978), p. 18; also Patricia Smith Churchland, "A Perspective on Mind-Brain Research", *Journal of Philosophy* LXXVIII, no. 4 (1980).

[8] See Eric Havelock, *The Literate Revolution in Greece and Its Cultural Consequences* (Princeton 1982), and Julian Jaynes, *op. cit.*, for interesting proposals about the cultural origins of mental concepts.

[9] See Karl Pribram, "Problems Concerning the Structure of Consciousness", in G. Globus, G. Maxwell and I. Savodnik, *Consciousness and the Brain* (New York 1976), and Edward Evarts, "Brain Mechanisms of Movement" *Scientific American*, (September 1979), 164–179.

[10] Newton, Natika, "Materialism, Metaphor and Mind", unpublished.

* * *

SOME REMARKS ON THE ROLE OF INDICATORS IN HISTORICAL INQUIRIES

Tadeusz Buksinski
University of Poznań

I. For the first time indicators were applied in a conscious and planned way in statistics. Then the concept was adopted by sociologists and psychologists of an empiristic bent. They conceived indicators as means for determining the empirical sense of terms which denote structures and factors unavailable to direct observation. The indicators were supposed to ensure social research, its empirical character.[1]

The aim of this paper is to determine the concept of indicator and to present the role of indicators in historical inquiries. Up to now attempts to reconstruct the process of historical inquiries based on the theory of indicators have not been made. The theory of history falls behind the other social and humanistic sciences in this respect.

According to my suggestion an indicator is every event from the occurrence of which one can infer the occurrence of another event. An indicator is any event A (a thing, a feature, a behavior, an utterance, etc.) which is empirical and related by regularity to another event B (a thing, a feature, a behavior, an utterance etc.). On the basis of their relationship and from the occurrence of indicating event A (referred to as indicator) we infer the occurrence of an indicated event B (referred to as indicatum). In science both the indicator and the indicatum belong to the class of empirical events. In science the role of the indicator may be played by any empirical event, both a directly observable and unobservable event.

To the most frequently applied indicators in history belong: 1) material objects and their features, e.g. a specific form of homesteads as an indicator of the material culture of the people living there. 2) written records (e.g. documents, chronicles) and their features (the information they include, the language in which they were written, the kind of materials on which they were written, etc.) 3) oral records and their features (sagas, legends, statements of witnesses etc.). 4) singular extra-source events such as the activities and customs of individuals and of groups and the specific feature of these activities and customs and their products. For the purpose of simplification all kinds of indicators will be referred to as an "event". We shall also avoid their time and space indexes. Such a broader concept of the indicator plays a useful role in considerations on the research procedure of a historian. The concept of the indicator is more universal than such concepts as 'source' and 'source information'. 'Lack of source'or-'lack of source information' and 'extra-source event' may also be indicators. The indicator makes it easier to work out a homogeneous research theory not only for history but also for all social sciences.

II. Substance relationships of various kinds may be developed between the indicator and the indicatum. The causes of events or their effects may be indicators. Events related to an indicatum functionally, genetically, structurally, etc. may be indicators. Events prior to an indicatum, simultaneous with or following it may be also indicators. But for the procedure of scientific research the most important meaning has the type of regularity or the degree of correlation between the indicator and the indicatum. On the basis of this criterion we devide indicators into exceptionless and probabilistic. The former are based on the following kinds of relationships: 1) relationships of exceptionless regularity ascertaining sufficient conditions, 2) relationships of exceptionless regularity ascertaining necessary conditions, 3) relationships of exceptionless regularity ascertaining both sufficient and necessary conditions.

In history the probabilistic indicators play a more significant role than the exceptionless

ones. The following indicators belong to them: 1) statistical indicators, 2) indicators presupposing general alternative regularities, 3) typological indicators. I will describe more accuratly solely the typological indicators. It seems that they are less known for theorists of science, even though they play an important role in researches.

General typological statements describe regularities which are neither exceptionless nor statistical. They ascertain relationships weaker than exceptionless, though in a sense stronger than statistical. They describe regularities of variable probability that may also be graded. The general obligatory principle is that the more the features of a definite situation (a thing, a person) are typical or model, the greater the probability of the occurrence of a typical indicatum. The more the given situation deviates from a model (type) the greater the probability of the occurrence of the modified indicatum.[2] Examples of typological statements: The more positive the evaluation an individual enjoys in a group the more he accepts the standards prevailing in the group. Of two fighting armies the more easily one army achieves the victory the greater its supremacy in numbers of soldiers and in arms and the better its command competence, and the higher the moral of soldiers.

A peculiar feature of typological statements is also the fact that unlike statistical statements they may refer also to situations which seldom emerge.

Cultural norms and social standards (the rules of behavior) are a subclass of typological statements. Cultural norms constitute a relatively stable pattern of human behavior. However, not every uniformed behavior pattern is a cultural standard, but only those which are caused by the influence of the external or internal normative constraints. As for the external pressure, it may be the pressure of the group we belong to or the institutions we work in, while to internal constraints belong moral principles. An indispensable condition to recognize an activity as attaining a cultural standard is the conviction of an individual that a given activity pattern is at least to some degree right. To some extent, the individual must internally accept a given way of activities and agree with it. To behavior patterns guided by cultural norms (standards) belong: a sumptuous life of the affluent European nobility; duels among the nobles in the case of an honour insult, etc.

Indicators based on cultural (social) norms have a probabilistic character. The way in which they are probabilistic is different from that of statistical indicators. The probability degree of the occurrence of the typical (model) indicatum is of a range nature in a sense and may be graded. Cultural (social) norms are not met by all and not always. The severer the punishment for failing to meet the ruling standards and the more the individuals accept their social enviroments, roles, statuses, and the stronger inner value system they posses, and the greater the extent to which social eviroments is typical ('normal'), then the greater the extent to which the individual activities are in accordance with ruling norms.[3]

The rules and principles of inferring from the historical sources have also a typological character. We will show it in the next chapter.

III. Indicators play the basic role in two research contexts: discovery and justification. First, indicators suggest hypotheses which may solve a posed problem. They indicate what hypotheses may be taken into account as possible and probable answers to a posed research question. Every statement in historical works is a hypothesis answering the question, which is chosen from many other possible answers to that question. Every statement may be confirmed and disconfirmed. A researcher cannot make haphazard or wild guesses; he offers and considers proposals for solving a posed problem on the basis of acquired knowledge important for a given issue. He particularly refers to the knowledge of detailed facts and to the knowledge of generalizations characteristic of an examined period and region. This knowledge performs the function of a set of indicators. It lays down the direction in which solutions to a posed problem should be sought. The better the historian, the more regularities and detailed facts of a given period he knows.

The process of discovering (constructing, propounding) hypotheses and the process of justifying them are closely related and dependent upon one onother. At the very moment of its

propounding a hypothesis is based on source-and extra-source indicators that support it. In the light of these indicators a researcher recognizes it as a possible and probable solution to a given problem. It may be generally assumed that the processes of heuristics are conclusion-drawing without the act of assertion: proposing from among all possible hypotheses concerning a given theme, the hypotheses which seem probable, even though their logical value and the degree of substantation have not been exactly determined. That is the stage of the process of research at which hypotheses are not accepted yet. They are taken to be propounded as for the sake of testing. Usually, there are many such hypotheses-suggested solutions to a given problem.

From the beginning the process of discovering is entangled in the process of challenging and eliminating other competitive hypotheses. Competitive hypotheses are all those probable solutions to a given problem which have more or less the same informative degree. Researchers try to determine the probability degree of each of the propounded hypotheses in order to accept the most probable. Eliminating competitive hypotheses in the process of justification takes place on the basis of analyzing the indicators known before the formulation of a given hypothesis and on the basis of those which will be discovered or determined in the course of justifying it. A very important role is played in that processes by typological principles of inferring from the historical sources. One of these presupposed principles may be formulated as follows: The better the author of a source is informed, the greater his interest in a given kind of facts, the more truthful and authentic the source is, then the greater the probability that the events described in the source did take place and that the unrecorded events did not occur. Assuming this kind of general statement the historians determine in a given situation the justifying force of a given indicator-event and as a result they estimate the confirmation degree of a given hypothesis. The followers of a hypothesis try to show that a given indicator quoted to support a hypothesis certainly occurred and that it confirms the hypothesis with highest probability (e.g. the source is highly reliable). The critics of a hypothesis try to show something contrary. Arguing, the oponents use the further range indicators, e.g. indicators of indicators.[4]

IV. It seems that on the basis of the analysis of the process of research one may attempt to compile the criteria according to which the confirmation degree of competitive hypotheses are estimated and compared. 1) The bigger the number of confirming indicators and the smaller the number of indicators invalidating a hypothesis, the better it is confirmed. 2) The new indicators discovered only after propounding a hypothesis confirm it to a greater degree than the old ones. 3) The more independent of one another the indicators are, the greater justifying value they possess. 4) The greater the variety of indicators, to a greater extent they justify hypotheses. The criterion of their variety is their attachment to different categories of events (things). 5) The greater the certainty of the occurrence of the justifying indicators (events) the better the hypothesis is justified (confirmed, disconfirmed). 6) The greater the probability relation between an indicator and an indicatum (a hypothesis) the better the hypothesis is justified. 7) The greater the degree to which an indicator challenges all but one of the competitive hypotheses, the greater its confirmation power. 8) The better (more accurate) the fact is known, the easier it is to determine its indicatory functions and to a greater extent it justifies the hypothesis.

I have mentioned solely the criteria, not having space to describe them in detail.

ENDNOTES

[1] Lazarsfeld, P. F., Rosenberg, M. (eds.), *The Language of Social Research* (Glencoe Illinois 1957). D. Lerner (ed.), *Evidence and Inference* (Illinois 1958). Bondon, R., Lazarsfeld, P. F. (eds.), *Le Vocabulaire des sciences sociales* (Paris 1972).

[2] Hempel, G., Oppenheim, P., *Der Typusbegriff im Lichte der neuen Logik* (Leiden 1936). Max Weber, *Wirtschaft und Gesellschaft* (Tübingen 1972) 5. Auflage.

[3] Nowak, Stefan, *Understanding and Prediction* (Dortrecht 1976).

[4] Buksinski, T., *Metodologiczne problemy uzasadniania wiedzy historycznej* [Methodological Problems of the Justification of Historical Knowledge], (Warszawa-Poznań 1982). The book presents the accurate analysis of many examples of justification of historical hypotheses.

* * *

METAPHYSICS, MIND, AND PHYSICALISM

Mark Migotti
Yale University

The aim of this paper is to criticize *physicalism* without compromising *naturalism*. By naturalism I understand the thesis that the mind is continuous with the rest of nature and that there is no deep ontological or epistemic divide between, on the one hand the mind and our understanding of mental phenomena, and on the other hand the rest of the world and our understanding of whatever phenomena it contains. By physicalism I understand the view that 'the physical' in some important sense has ontological primacy over 'the mental'. I shall attempt to show both that it is useful to distinguish physicalism from naturalism and that physicalism entails some implausible claims about the scope and nature of physical science. I shall present my case via a critique of K. V. Wilkes' recent and sophisticated arguments for a physicalistic philosophy of mind found in her 1978 book *Physicalism*. (Hereafter referred to in the text as *W*.)

I

Wilkes takes the central issue addressed by physicalism to be that of "whether the vocabulary and theories of physical sciences are adequate to describe, predict and explain the purposive behaviour of humans and animals. . . ." (*W*, p. 10.) According to Wilkes, the defining characteristic of the vocabulary of physical sciences is extensionality. The language of theories from the physical sciences, that is to say, consists exclusively of sentences that admit of the substitutability of co-referring expressions *salva veritate*. The central claim of physicalism, then, becomes that human behaviour can eventually be explained with the use of extensionally formulatable theories alone. Appearances to the contrary, intensional operators such as 'believes', 'hopes', 'fears' and so on are contextually eliminable for purposes of explaining and predicting human action.

Wilkes acknowledges that, *pace* radical behaviourism, one cannot successfully eliminate intensional locutions from psychological theory by simply refusing to touch them. In the wake of arguments from Dennett and others Wilkes is convinced that "behaviourism fails to explain or predict . . . some of the very simplest behaviour of mice and men." (*W*, p. 18 and ff for references to the literature on behaviourism.) Trying to account for why people do the particular things they do without reference to what they desire to accomplish and to what they believe, hope or fear about their situations—in short, without reference to how they represent the world to themselves—is, it seems, akin to entering a tennis match armed with a ping-pong paddle; one's equipment just isn't up to the job.

Wilkes' response to the demise of behaviourism is, briefly, to try to turn to physicalist ends the explanatory power afforded by the use of propositional attitudes and other intensional concepts in psychological theory. In other words, *contra* behaviourism, intensional sentence forming operators are to be permitted in the course of explaining intelligent behaviour, but as per physicalism, intensionally loaded explanations are to be officially endorsed only if they are translatable without loss of explanatory power into explanations employing only an extensionally pure language. Functionalism, Wilkes argues, is a methodology expressly designed to allow the science of psychology to benefit both from the epistemic advantages of an intensionally laden theoretical vocabulary—since without it not enough explaining gets done—and from the ontological advantages (as they are perceived) of physicalism—since the alternative is assumed to be some untenable form of dualism.

Here, roughly and telegrammatically, is an outline of Wilkes's functionalist methodology. First, one treats psychological theories as trying to account for various mental capacities by positing a number of pre-conditions which (it is conjectured) must be met if the capacities are to be used as they are. In other words, one regards the *explananda* for which psychology tries to account as a number of operationally defined systems consisting of systematically interrelated and organized parts. The leading question then is: How does the system carry out its characteristic functions? In short, How does it do what it does? The sort of answer one is looking for will be of the form 'it manages to do G because it does P, Q, and R, and doing P, Q, and R in a certain way just *is* doing G. Thus, to use a favourite example of Wilkes', a washing machine performs its characteristic G, washing clothes, by a) soaking, b) soaping, c) agitating and so forth. Now each of the functional prerequisites of a system can in turn be the object of another functional explanation. How, for example, does the machine soak the clothes? Well, it a'. . . ., b'. . . ., c'. . . . etc. . Eventually one arrives at a level of explanation which counts as basic. 'Base level' as Wilkes points out, is equivalent to "what scientists are currently calling the 'base level'". (*W*, p. 64.) The scientists who decide what base level is are, of course, physicists. (see *W*, p. 30.)

It is perhaps now becoming clearer how Wilkes expects functionalism to provide the physicalist psychologist or philosopher with a cake which is simultaneously eatable and havable. Intensionally phrased explanations are to be encouraged as long as there is hope that ontological commitment to the objects of intensional discourse can be 'discharged' at the level of neurophysiology. Less and less of the language with which psychological abilities are characterized will be required as these abilities are explained as the product of a number of interrelated sub-abilities. Psychological theory will link up with neurological theory as the unit functions of the former discipline are broken down into huge, complex conjunction of specialized neurological sub-functions, culminating in the extensionally describable function of individual neurons which is simply to fire when a certain afferent energy threshold has been reached. As Wilkes puts it, ". . . the greater the detail of the research, the more inflexible will be the functions that are performed by micro-structures or micro-processes of the brain; intensional descriptions are needed only when functions are complex and flexible." (*W*, p. 65.)

Borrowing a phrase of Kant's, I feel that this apparently tolerant functionalistic physicalism hides a significant "danger to our thinking self". (*K.d.r.V.*, A/383.) My principle fear is that the model of explanation of mental phenomena Wilkes espouses intimates a lot more for itself than it delivers. It does this by trading on an ambiguity between a strong, controversial interpretation (which I would call physicalism proper) and a mild, uncontentious interpretation (which I would call naturalism) of her central thesis. In the next section I try to show how this ambiguity can be focussed by examining the role of the physical sciences in our knowledge of the world as a whole.

II

Wilkes' attitude towards natural science is typified by the following two quotations:

> On the physical side of the mind-body relation . . . there seems no bar in principle to the suggestion that the physical sciences which study the brain and human activity can have, or devise a vocabulary adequate for the identification of any physical state, event or process. (W, p. 2.)
>
> . . . whatever one's criteria for a good theory may be, there seems no reason why (the physical sciences devoted to the study of the brain and the central nervous system) should not eventually meet them. (W, p. 31.)

Given that for Wilkes "the scope of 'the physical' is the physical scientist's business." (*W*, p. 2.) the first claim strikes me as self-authenticating in a way which renders it vacuous. It is as if a 'magicalist' engaged in a hypothesized controversy over the 'magical-mundane' problem

were to argue as follows: "'the magical' picks out all and only items, processes, concepts, laws, hypotheses, theories or theoretical postulates used essentially by magicians. There seems, therefore, no bar in principle to the suggestion that qualified magicians studying the universe and the spell books can have or devise a vocabulary adequate for the identification of any magically caused events, states or processes." Of course (given these definitions) the physical scientist and the magician can pick out instances of the physical and the magical respectively, they are picking out just what they posited in!

The second passage, I presume, contains a likely response to such an attempted *reductio ad absurdum*. The physical sciences, it might be said, are capable not just of constructing any old theories about the world, but more importantly of constructing *good* theories about the world.

But even if it is accepted that good theories are (*inter alia*) those that permit successful prediction and control of the phenomena allegedly explained and that physics and chemistry are exemplary producers of good theories, it does not follow that theories which use an intensional vocabulary (apparently) not remotely connected with the extensional language of physics and chemistry cannot therefore be good theories; it does not follow, that is, unless one makes the very strong claim that *really* physics and chemistry are the *only* producers of good theories. To make this latter claim is to place a virtually a priori faith in the foundational position of physics *vis a vis* the rest of theoretical knowledge. It is also, in effect, to deny that any of the social sciences are capable of constructing good theories, unless the rider 'a theory which *might* one day be translated into extensional, physicalist language can still count as a good theory' is construed so broadly as to let in anything one likes. Finally, the claim also implies that probabilistically framed theories are dubious since 'the probability of x/y' is (on many accounts) an intensional sentence forming operator.[1]

If one endorses these consequences of strong physicalism, one is committed to the view that science properly so-called, i.e. physical science, has nothing at all to say about a vast range of phenomena, including most of the more interesting aspects of recorded human history. If, as seems to me immensely more attractive, one regards the whole of science as trying to explain the whole of experience, then it might well make sense to use different models and vocabularies for different sectors of experience as required. The best theories about the causes of wars use mainly intensional socio-economic terms while the best theories about the causes of plagues generally get by with extensional bio-medical terms. Notice, though, that the difference in vocabularies does not prevent theories from the two domains from interacting; wars can be among the causes of plague and plagues can be among the causes of war. (I take it as obvious that wars can not be adequately described in extensional terms alone.) Furthermore, if what constitutes a good theory is itself a theoretical issue—and if one acknowledges a discipline called 'epistemology', this is precisely its field—then one cannot (or should not) *dogmatically* privilege extensional, physicalistic vocabularies over intensional, mentalistic ones without at least suffering pangs of philosophical conscience. One should, that is, have an argument to show that the epistemic benefits of extensional purity outweigh the costs of sacrificing a large body of apparently serviceable knowledge about social behaviour which seems inextricably bound up in the intensional idiom.

Wilkes has no such argument to the best of my knowledge and I conclude that her brand of physicalism fails to account for some of the most interesting features of the mind, preeminently features expressed in social life, and that it furthermore rests upon a misguided hierarchical vision of science with physics as the solid foundation. Only a dogged and misconceived 'philosophical faith' (to echo William James) in physical science could lead one to hold that explanations in terms of propositional attitudes and other intensional descriptions of mental life can be countenanced only as temporary stand-ins for the pushes and pulls and charges and discharges that really underlie the whole show.

ENDNOTE

[1] I owe this point to Hugh Mellor whom I thank for helpful comments on the matters discussed in this paper.

REFERENCES

Kant, Immanuel *Kritik der reinen Vernunft*, English trans. Max Muller (New York 1966)
Wilkes, K. V., *Physicalism* (London 1978)

* * *

INSTRUMENTALISM IN THE PHILOSOPHY OF MIND

Gary Fuller
Central Michigan Unversity

A central concern in philosophy is the relation between common sense and science. Philosophy of mind, of course, has shared, and continues to share, this concern. Here the issue is that of the relation between our commonsense psychological notions of belief, desire, etc. (and the related notions of personhood, moral agency, and responsibility) and the rising sciences of the day: traditionally, Newtonian mechanics; some decades ago, radical behaviorism; and now, cognitive psychology and neurophysiology. The general question has remained constant: does science clash with our folk-psychological notions, or can it exist in harmony with them? Today, there are both optimistic thinkers, such as Jerry Fodor,[1] who are confident that an adequate scientific psychology will turn out to be quite similar to folk psychology, as well as pessimists, such as Paul Churchland,[2] who believe that the developing psychological and brain sciences will prove to be totally incompatible with folk psychology and that folk psychology should therefore be abandoned altogether.

One way in which a philosopher might try to protect folk psychology against possible threats from these advancing sciences is by adopting an instrumentalist interpretation of folk psychology according to which folk-psychological terms, such as 'belief' and 'desire' do not refer but rather are to be regarded as mere calculation devices. Daniel Dennett, one of the most influentional philosophers of mind during the last fifteen years, is well known for his espousal of such a position. The aim of this paper is to examine instrumentalism through the work of Dennett and to raise the question of whether instrumentalism vindicates folk psychology. I shall do three things: (I) clarify the notion of instrumentalism; (II) give a brief account of Dennett's instrumentalist views; and (III) argue that at least Dennett's earlier strong instrumentalism fails to vindicate folk psychology.

I

Let us begin by giving some clarification of the notion of instrumentalism. What does it mean to be an instrumentalist with respect to folk psychology?

It will be easiest to spell out what instrumentalism involves if we make an assumption which is now widely held: the assumption that folk psychology is a *theory* of behavior implicit in our explanations and predictions of behavior in terms of beliefs, desires, and so on. On this assumption the mental terms of the theory, e.g. 'belief' and 'desire', are to be construed as theoretical terms which at least ostensibly refer to state types and/or tokens.

Having made this assumption, we can give an initial characterization of instrumentalism by contrasting it with realism, as follows:

(1) Realism: the mental terms of folk psychology refer to state types and/or tokens and its sentences have a truth value.

(2) Instrumentalism: the mental terms of the theory do not refer and its sentences have no truth value. Rather, such terms and sentences are calculation devices for the purpose of predicting behavior.

Notice that according to our characterization so far instrumentalism is incompatible with identity theories of both the mental type/physical type as well as the mental token/physical token varieties.

This initial account of instrumentalism, however, can be refined by relabeling (2) 'strong instrumentalism' and distinguishing it from weak instrumentalism, which can be characterized as follows:
(3) Weak instrumentalism: the mental terms of folk psychology do not refer; nevertheless, the sentences of the theory do have a truth value.

II

Over a number of years Daniel Dennett has articulated and defended an instrumentalist approach to folk psychology. Dennett's instrumentalism has not remained static; indeed, a good case can be made for claiming that his instrumentalism has undergone a major change.[3] Earlier on, especially in *Brainstorms*,[4] Dennett held a strong-instrumentalist position in which the sentences of folk psychology were construed as having no truth value. More recently, and here the key article is "Three Kinds of Intentional Psychology,"[5] he has adopted a weak instrumentalism according to which folk-psychological states, such as beliefs, are not genuine posits (*illata*) but mere calculation bound entities (*abstracta*) analogous to centers of gravity. On the new view then, the term 'belief' does not refer, but belief *sentences*, just like sentences "about" centers of gravity, do have a truth value.

Putting these details to one side, we need to ask the general question of why Dennett has favored instrumentalism at all, rather than opted for a realist position. Certainly, one central consideration for Dennett has been that folk psychology (or Dennett's refinement of it, Intentional Systems Theory), if interpreted realistically, might well prove to be false. Among other things, folk psychology contains a rationality idealization and is therefore unrealizable, and a realistic construal of folk psychology might seem to require inner "representations", about which Dennett has often been sceptical, to say the least.

This consideration taken by itself is obviously not enough. If folk psychology taken realistically is literally false, then why not scrap it altogether? There are a number of reasons why Dennett wants to preserve folk psychology. To begin with, Dennett believes that it does have a role to play in scientific psychology: it can function, for example, as a useful starting point in top-down strategies. In addition, and perhaps more important, folk psychology is in some sense indispensable to us. The second major consideration for Dennett in favor of adopting instrumentalism, then, is that it will enable us to preserve folk psychology. Instrumentalism will in a sense vindicate folk psychology.

III

What, if anything, is wrong with an instrumentalist approach to folk psychology? In this section I shall restrict myself to criticizing the strong instrumentalism of Dennett's early view, indeed, to exploring just one strand of criticism of the early view. I want to question whether strong instrumentalism, whatever its other merits and weaknesses, really does vindicate folk psychology.

Does the strong instrumentalist construal of folk psychology *vindicate* our folk notions of belief and desire and the related notions of personhood and responsibility?

Now what does 'vindicate' mean here? One sense in which a claim or theory, say in the face of a sceptical attack, can be said to be vindicated is by being shown to be true after all. But this sense, of course, is not the one we want here. 'Vindicate our folk-psychological notions' must mean here: 'show that we have good reason to continue to use these notions.' Taking 'vindicate' in this way, then, we might at first glance be tempted to argue that strong instrumentalism does vindicate our folk notions: surely the usefulness of these notions as calculation devices provides ample reason for continuing to employ them.

The trouble with this is that our folk-psychological notions are not mere calculation devices.

Indeed, as a number of philosophers have correctly argued, a strong-instrumentalist account of our idioms of belief, desire, and personhood will yield concepts different from our actual ones. Our actual concepts are *realistic*: facts about the inner states of a system (and about their causal relations to each other and to inputs and outputs), are relevant to the application of these concepts in a way that they are not to the application of the corresponding strong-instrumentalist concepts. Contrast, for example, the strong-instrumentalist concept of a "person" (or "mind") according to which, following Dennett, a system counts as a "person" just in case its behavior is "reliably and voluminously predictable"[6] via the instrumentalist rules of the intentional strategy, with our actual concept of a person. The two are clearly different, since:
(1) We can describe systems which satisfy the strong-instrumentalist but not the actual concept of a person. Stephen Stich has developed Ned Block's example of a chess-playing computer whose moves are generally predictable using the intentional strategy, but which has no internal representations of rules or goals.

> Rather, this computer's memory contains an *enormous* multiple branching tree representation of every possible chess game up to, say, 100 moves in length. ...if we were to run across (what appeared to be) a person whose conversations, chess playing, and other behaviors were controlled by an enormous preprogrammed branching list of what to do when, I think our intuition would rebel at saying that the "person" believed that I was about to attack with my queen—or indeed that he believed anything else!'[7]

(2) It is also easy to think of examples in the converse direction, examples of systems which satisfy the actual but not the strong-instrumentalist concept of a person. My colleague, Paul Yu, came up with the following case. Suppose that, unbeknownst to us, Martians so arrange matters that Jones's brain states become erratic whenever any attempt is made to predict his behavior, although he is perfectly normal at other times. Then, although Jones would count as a person on the actual construal of 'person' he would fail to be one on the instrumentalist construal.

Our actual folk-psychological concepts, then, are different from their strong-instrumentalist counterparts. Now if Dennett were interested in folk psychology only for the role that it can play in scientific psychology, e.g. (as we have seen) in developing top-down strategies, then perhaps this lack of equivalence need not worry him. But if, as he seems to be, he is concerned with vindicating our ordinary psychological notions and practices, then he *should* be worried: for how can the introduction of instrumentalist substitutes, however scientifically legitimate, vindicate our actual *realistic* concepts of desire, belief, and personhood?

To answer this last question we might try the following line. We use our folk-psychological concepts to draw certain distinctions: that between persons and non-persons, that between actions for which a person can be held accountable and those for which he cannot, and so on. Perhaps, as Dennett seems to think, these commonsense concepts are flawed in certain ways and should therefore be replaced by new unflawed concepts—in Dennett's opinion, by the new strong-instrumentalist substitutes. If the new concepts draw roughly the same distinctions as the old, then we can say that the new concepts have vindicated the old distinctions, and, indirectly, the old concepts. Taking this line of thought, we can hold, then, that the instrumentalist substitutes will vindicate our actual folk concepts just in case they agree reasonably closely in their (possible) extensions.

How close is reasonably close? Examples (1) and (2) above have already shown that the strong-instrumentalist substitutes are not equivalent to our actual folk-psychological concepts. Still, it might be thought here that the situations described in (1) and (2) are very unlikely to occur and that the instrumentalist and the actual concepts will turn out to be at least co-extensive in the actual world. Surely, this is enough for vindication.

Unfortunately (for the strong instrumentalist) we can think of cases which are less unlikely to occur but in which there is also a divergence. Consider the following example:
(3) Jones is pointing a gun at Smith and has every intention of pulling the trigger and killing him. He subsequently pulls the trigger and Smith dies. Nevertheless, what caused Jones's pul-

ling of the trigger was not his intention, or any other mental event, but rather some fluke brain event totally unrelated to any of Jones's mental states.
Although this example needs to be described in much greater detail, a good case can be made for claiming that we (using our ordinary folk concepts) would not want to hold Jones responsible for Smith's death whereas the strong instrumentalist would.

In conclusion, then, I have serious doubts about whether Dennett's earlier strong instrumentalism can vindicate our ordinary folk notions of belief, desire, and personhood. Indeed, I suspect that Dennett had these doubts as well and that they were an important motivating factor in the change from his earlier view to his new weakinstrumentalist view. Whether the new view also runs into similar problems about vindication is something which must await further exploration.*

ENDNOTES

1 See his *The Language of Thought* (New York 1975).
2 See his "Eliminative Materialism and Propositional Attitudes", *Journal of Philosophy*, Vol.78 (1981), pp. 67–90.
3 Paul Yu and I have argued this case in "A Critique of Dennett", (forthcoming in *Synthese*).
4 Dennett, Daniel, *Brainstorms* (Cambridge, Mass. 1978).
5 In R. Healey (ed.), *Reduction, Time and Reality* (Cambridge, England 1981), pp. 37–61.
6 "True Believers...", in: A. Heath (ed.), *Scientific Explanation* (Oxford 1981), p. 55.
7 Stich, Stephen, *From Folk Psyhology to Cognitive Science* (Cambridge, Mass. 1983), pp. 244–5.

* I am heavily indebted to Paul Yu: this paper grew out of many long and enjoyable discussions with him on Dennett over the last year. Thanks also to Regine Imbsweiler, Ausonio Marras, and Salma Saab for helpful comments.

* * *

CAN HUMAN CONCEPTUAL SCHEMES BE STUDIED EMPIRICALLY? SOME CAUTIONARY REMARKS

Sharon Lee Armstrong[1]
Wesleyan University

The tradition in philosophy and psychology has been to assume that common words such as 'bird' are cover-labels for mental categories that are themselves bundles of simpler categories called features, attributes or properties. There are two major subtypes of a feature theory approach to mental categories: the *definitional* view, and the *prototype* view. On the definitional variant, a smallish set of simple properties are individually necessary and severally sufficient to pick out all and only, say, the birds, from everything else. While meritorious for its simplicity and precision, this view has proved recalcitrant to work out in the required detail. It's rarely the case that all and only the class members can be picked out in virtue of sufficient lists of necessary (let alone elemental) properties.

However, there is another class of feature descriptions that relinquishes the necessary and sufficient requirement. This is the family resemblance view first alluded to by Wittgenstein (*PI*) although he might be surprised at some of its recent guises. He took as an important example the concept *game*, held together by a variety of gamey attributes, only some of which are instantiated by any one game.

Psychologists have adapted this view and brought it into psychology through a series of compelling experimental demonstrations as the 'prototype' theory. The idea here is that there are distinctions among the properties themselves, some recognized as privileged relative to a category. Prototypical members have all or most of the privileged properties of a category. Marginal members have only one or a few. So for *bird,* we have robin as a prototypical member exhibiting the most privileged properties (flies, wings, feather-bearing, egg-laying), whereas penguin exhibits fewer of these properties and is considered marginal by many. This approach also offers to give an account, in terms of various measures of feature overlap and/or organization, the apparent fact that membership in a category is graded; for example, why a robin seems birdier than a penguin.

However, there are grounds for caution before embracing a particular interpretation of these findings. It is not notably easier to find the prototypic features of a concept than to find the necessary and sufficient ones. Prototype theories render the description of reasoning with words far more difficult and the understanding of compositional (phrase and sentence) meaning altogether hopeless. Thus it seems surprising that psychologists have been pleased, rather than depressed, by experimental findings that support a prototype theory. These problems provide some impetus to reconsider the empirical support.

Armstrong, Gleitman & Gleitman demonstrate serious empirical challenges and give extensive methodological criticisms of the empirical base for the prototype approach. I shall just quickly review some of their findings here.

Following upon Rosch, subjects were asked to indicate how good an example various instances of categories are by using a rating scale from 1 (good) to 7 (poor). It turns out that people will say that apples are very good examples of *fruit* and deserve a high rating while figs and olives are poor examples and deserve low ratings. When subjects were asked to judge the veracity of statements about category membership, it turns out that they respond faster to "AN APPLE IS A FRUIT" than to "A FIG IS A FRUIT" with word frequency controlled across the list. Rosch and her colleagues have interpreted these findings as evidence that membership is graded, rather than all-or-none in everyday categories. But if certain concepts are nondefinitional because of graded responses, that must be because if the concepts were

definitional, graded responses would not be achieved. But this remains to be shown by testing some definitional concepts.

Are there definitional concepts? Of course. For example, consider the superordinate concept *odd number*. It is precisely defined as "an integer not divisible by two without remainder." No integer seems to sit on the fence, undecided as to whether it is quite even or perhaps a bit odd. No odd number seems odder than any other. A, G & G found that subjects gave graded responses to the definitional concepts *odd number, even number, plane geometry figure*, and *female* in both the exemplariness-rating and verification tasks!

These results suggest that we are back at square one in discovering the structure of everyday concepts *experimentally*: Certain techniques widely used to elicit the structure of such categories are flawed and thus the structure of those concepts studied remains unknown.

Perhaps a more direct approach to the study of the internal structure of everyday concepts would be more profitable. Instead of focusing on the kind of organization of features (definitional vs. prototypic) for a concept, focus instead on the features themselves. In psychology, the most direct approach has been taken by Rosch & Mervis. Subjects were presented with various exemplars from a number of superordinate categories (e.g., chair, sofa, bed, from the category *furniture*) and asked to list all the attributes that are "common to and characteristic of" each item. Their rationale was straightforward: If there is a set of necessary and sufficient attributes that defines, say, *furniture*, then every item that falls under concept necessarily has all the required attributes and these should emerge in the data collected.

In fact, Rosch & Mervis found that "very few" (sometimes no) attributes were listed for all the exemplars of a category. But it is by no means clear that the subjects could really comply with the instructions to come up with the appropriate features that describe a given category. And what does one make of the features that subjects do produce?

One problem concerns the suppression of features. A subject would not list *all* of the attributes of a given term even if he knew them. Some of the reasons have to do with lexical redundancy rules: Most subjects won't mention *living thing* let alone *physical object* for penguin. The features of the superordinate are simply presumed to apply to the items that fall under it. For related reasons, people tend to tell you what they think you need to know, suppressing the obvious. A standard dictionary defines *zebrula* as a "cross between a zebra and a horse"; but no dictionary (or subject) would ever define a *horse* as a "cross between a horse and a horse". Thus, not only will particular attributes vary from item to item within a superordinate category but level of response will likely vary as well.

An even more troublesome problem is whether the subjects could access the features and express them in words. Some empirical basis for this worry comes from pilot data (with Komatsu) partially replicating the Rosch & Mervis study. On the face of it, subjects' responses to *grapefruit* and *tractor* showed considerable variation. For *grapefruit,* subjects responded "sweet", "sour", and "not as sweet as an orange". For *tractor*, subjects responded "has four wheels", "has three wheels", and "has two wheels (very unstable)". Yet the variability may only be apparent. Notice that none of them said how many wheels a grapefruit had nor how sweet a tractor was. The subjects seem to share some common conceptions of the categories but perhaps were unable to come up with the right level of description—e.g., "bewheeled" and "sweet/sour dimension". Even if categories are describable in terms of some featural vocabulary, it will be difficult to expose this directly by limited inquiry.

Leaving aside questions of the interpretation of subjects' responses, there are other issues regarding variability of responses. Those feature lists that do emerge from psychological studies of concepts are obtained by pooling data across many subjects. This move presumes that concepts are stable across individuals in the relevant way. Never subjected to empirical test, this assumption is suspect. Suppose each subject's mental representation for *fruit* were definitional, but each subject's definition differed ever so slightly. Pooling data across subjects would lead the researcher to conclude that the concept *fruit* is a family-resemblance concept. But is it? It depends on whether your interest is in the linguistic community or in the individual. Current research methods would prejudice the results to be

descriptive of the community even though the researcher's intention is to study the individual (collectively).

Problems of the sort just mentioned lead me to consider, in general, the efficacy of the empirical study of conceptual schemes as it is practiced now in psychology. It will be useful to contrast the philosophical and psychological approaches to the study of concepts.

The philosophical study of concepts is best represented by the approach called 'conceptual analysis'. It is characterized by extensive and penetrating probing of one's intuitions about the character of a particular concept. It involves the pitting of intuitions against intuitions, the attempt to generate counterexamples to any initial hypotheses, the attempt to devise questions and contexts, to reveal *the* nature of the concept under investigations, i.e., what *ought* to be believed by truth-seekers. Thus, the philosopher relys primarily on himself as an authority, but may draw upon other philosophers' beliefs subjected to the same rigorous probing.

In contrast, the psychologist's intent (usually – although some psychologists are confused about what their mission is) is to study the 'collective individual'; He would like to determine the nature of a conceptual scheme found in *most* people, which identifies this enterprise as empirical. The representativeness of a sample studied being an issue, many subjects (sometimes hundreds) are polled, responses tabulated, and statistical conclusions drawn. Subjects' responses are not typically analyzed vis a vis one another but simply tallied. In fact, there is a reluctance on the part of many researchers to provide much in the way of interpretations adhering to the dictum "let the data speak for themselves". The fear evoking this policy is that the investigator might inject too much of his own perspective into an interpretation and that conclusions drawn from tabulated subjects' responses are, in contrast, 'bias free'.

Bent on obtaining large numbers of subjects, psychologists have (due to oversight or the constraints of time and funding) limited both the range and depth of questioning to each subject (introducing bias nonetheless). Thus, the character of the data collected will be importantly different from that of the philosopher. The question arises, "How can we be sure that a severely limited sample of any particular subject's intuitions (however many times multiplied), would reveal the richness and complexity of human conceptual life that we know to exist?" I think that the examples given in the discussions above indicate that we are unlikely to.

To the extent that one's interests are directed to the collective individual, one *must* question a reasonable number of subjects but hundreds are not usually necessary or even better. (Repeated-measures designs offer some statistical advantage in terms of reducing one source of variation.) However, any attempt to study conceptual schemes empirically that can hope to avoid the problems mentioned above will also have to adapt the extensive probing techniques of philosophy. This will necessitate limiting the inquiry to a few conceptual domains at a time and studying them in depth rather than limiting the questioning to each subject. This is just as well since it is unlikely, that discoveries concerning the various important conceptual domains will reveal that any of them are organized as simple feature structures. Rather, in each domain, the units, their patterning, the principles that organize them, their development, their environmental dependence are all likely to be different and complex, rewarding serious study.

REFERENCES

Armstrong, S. L., Gleitman, L. R., & Gleitman, H., "What Some Concepts Might Not Be", *Cognition* Vol. 13 (1983), pp. 263–308.

Rosch, E., "On the Internal Structure of Perceptual and Semantic Categories", in T. E. Moore (ed.), *Cognitive Development and the Acquisition of Language* (New York 1973), pp. 111–144.

Rosch, E. & Mervis, C. B., "Family Resemblances: Studies in the Internal Structure of Categories", *Cognitive Psychology* Vol. 7 (1975), pp. 573–605.

ENDNOTE

[1] I am indepted to Lila R. Gleitman and Henry Gleitman for many ideas and discussions of issues presented here.

* * *

DE-IDEALIZING THE MIND-BODY PROBLEM

John Blackmore
Kirchberg am Wechsel

I

If what is often called the mind developed during the course of evolution, then we may be dealing primarily with an historical problem which could vary from species to species and even from individual to individual, that is, the relation between mind and matter may be contingent rather than necessary.

In other words, philosophers may be unable to solve the problem merely by switching categories around or by "eliminating linguistic confusion". It may require a knowledge not only of physics, chemistry, biology, physiology, medicine, and psychology but of the history of human and animal evolution as well.

Unfortunately, however, many philosophers may find it difficult to understand how most scientists think, especially historical and experimental scientists. What most philosophers understand by law-like claims (often any unqualified statement)[1] is drastically different from the few actual claims which are widely acknowledged as laws by significant numbers of scientists themselves (idealized universal claims which describe how idealized types of variables behave under idealized types of limited-variable conditions). Real scientific laws such as those of Galileo,[2] Boyle,[3] and Charles,[4] represent preliminary stages of understanding which while true under idealized conditions tend to give way in practical terms to laws which take more variables into account, that is, until a level is reached where there are too many important variables to be handled by laws, such as is largely the case in biology, geology, and the social sciences.

To the extent philosophers understand science it is often theoretical physics which serves as a model (particularly work done during the first three decades of this dying century). This is partly because "explanation" can often be idealized as deduction of a factual or verifiable claim from what philosophers imagine is a scientific law (an approach not utterly unrelated to traditional syllogistic reasoning). Regrettably, theoretical physics is not typical of the sciences (and hopefully dubious processes such as "renormalization" will never be typical), few claims which philosophers call laws are really laws, and real explanation normally bears little resemblance to mere formal deduction.

II

If philosophers wish to understand what the scientists who are really making progress toward resolving the mind-matter problem are doing, then it could be wise *to learn how to think historically*, that is, how to *de-idealize* one's thinking and understanding. Experimental scientists who have studied biology and physiology in a successful manner over the last two centuries such as Claude Bernard, Louis Pasteur, Gregor Mendel, Charles Darwin, and in medicine Joseph Lister, Robert Koch, and numerous others while they were able to draw important generalizations from their work and use those generalizations to guide subsequent work nevertheless had a marked ability to *de-idealize* problems. They were able, first, to *de-typify* the phenomena to be studied, that is, concentrate on particulars as particulars, second, *de-generalize*, that is, keep within the evidence and qualify their understanding of results at every important stage, third, *de-formalize*, that is, avoid unnecessary mathematics, abstractions,

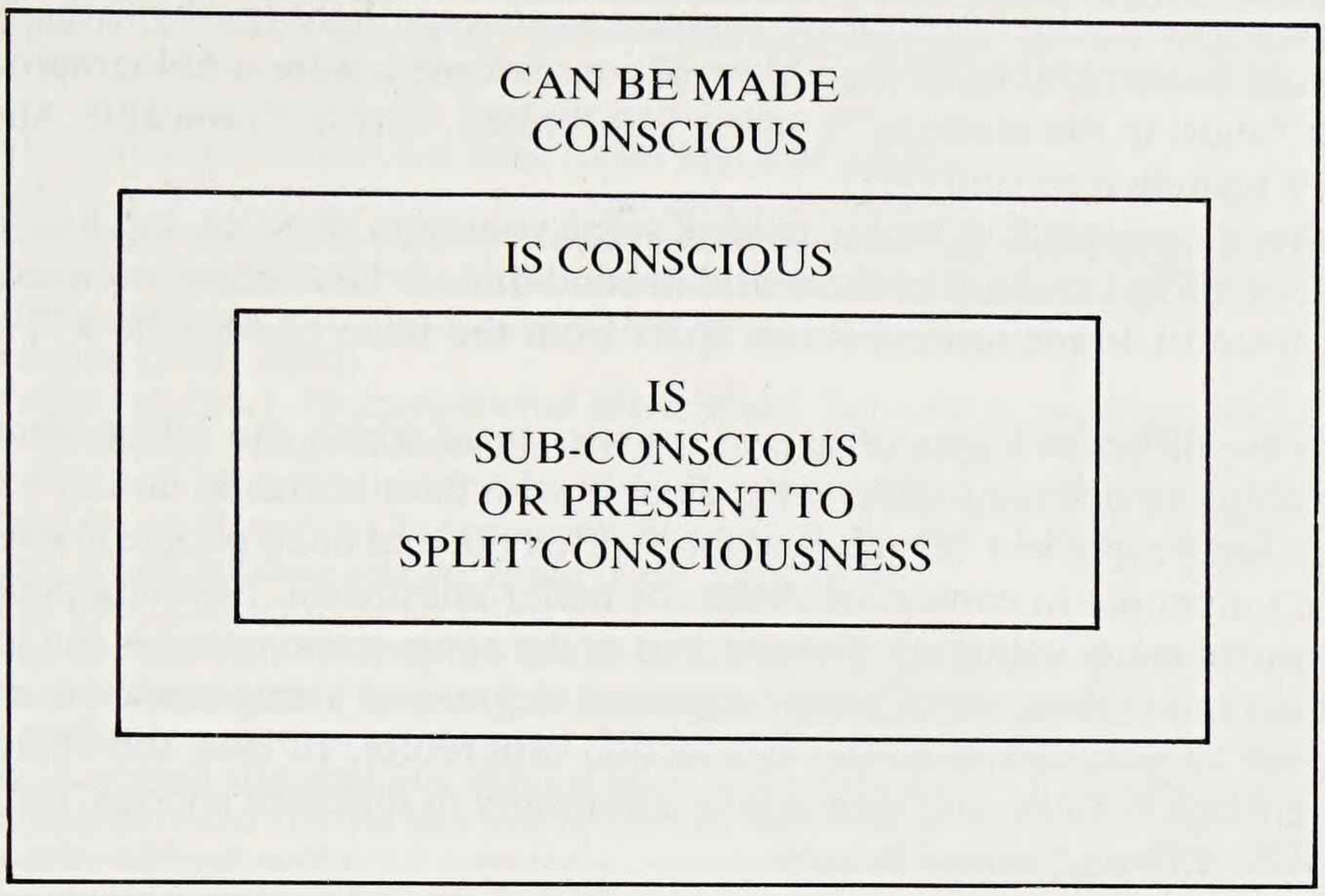
and technical jargon, fourth, *de-fictionalize*, illustrate and support arguments by real rather than fictional or imagined examples, fifth, *de-conventionalize*, base assumptions and arguments on facts rather than mere rules, definitions, or conventions, and sixth, *de-dogmatize*, hold no beliefs more strongly than the weight of evidence supports.[5] No one is perfect, and even great scientists can slip, but the above characteristics were generally present in their investigations.

There is no shortage of people inclined to jump to conclusions. Almost every important theory or discovery had its precursors. Copernicus had his Aristarchus, Dalton his Democritus, Newton his Hooke, and Darwin had more than can easily be listed, including his grandfather.[6] It is massive factual support and careful detailed study which can select the winner from a thousand hypotheses. Darwin's thesis wasn't new, but no one else took the trouble to study individual species and their history in such an extensive way to support that thesis. The mind-body problem is not short of theories or hypotheses or of switching the boundaries between mind and matter or abolishing them altogether. What is missing among philosophers is the patience and scientific understanding to study and collate the relevant data in a reasonably exhaustive manner. What is missing among scientists is of course lack of philosophical understanding. What is usually missing in both is a careful grasp of the earlier history of the relevant sciences. But the Darwins of the world can combine the characteristics needed.

III

Since Descartes the key to the mind-body problem has generally been understood to be consciousness. It is this issue which requires the most de-idealization if substantial progress or an experimental breakthrough is to be accomplished. Those philosophers who treat consciousness as an all-inclusive category such as Hegel[7] and Husserl[8] or as non-existent such as some early behaviorists not only share the same position in practical terms (i.e. an ability to ignore consciousness) but appear to be contradicted not only by a great deal of experimental work in biology, physiology, psychology, and medicine but in classical terms by Oswald Külpe and his Würzburg colleagues on the eve of the first world war.[9] They apparently discovered not only imageless thoughts (thereby helping to refute several old arguments by Berkeley and Hume against representational realism) but also that much or most imageless thinking is also unconscious. But if some thinking is conscious and some not, then neither an idealistic theory that everything is conscious or a behavioristic one that nothing is conscious can be sound. Consciousness is presumably real but has a limited and varying scope, that is, among those living creatures who are conscious.

Indeed, it is apparently a necessary condition for experimental study that what is studied be both conceivable as possibly real and limited such that one can at least conceive of or allow for a contrary, something to compare what is being studied against. If everything were conscious, how could we study it? What would we be studying which could distinguish it from what we are not studying? To be without a contrary removes our ability to understand it in a comparative way, and without at least some comparison can anything be said to be understood? In short, the scientific capacity to study consciousness already presupposes both the conceivable reality and the limited scope or range of consciousness. Without those two factors no scientific study seems possible. The following diagram could be helpful. Some things presumably are conscious (at least to some living creatures), some things which are not conscious can be made conscious, and presumably some things which are not conscious to anyone or any animal can never be made conscious to them, though depending on which epistemological assumptions are reliable, perhaps we can still understand or at least refer to much that is beyond the possible scope of consciousness, at least let us hope so. While the diagram below may seem to have general or universal application, in fact different creatures at different times and under different conditions may have their own diagrams as it were which possess different ranges of inclusiveness.



Different philosophers of course define "reality" in different ways. Philosophers sometimes link it to perception, others to consciousness, and still others to Platonic Forms or universals. Most scientists and practical people, however, define reality as what can make a causal difference, that is, as a power or force, as if Hume's arguments were either false or irrelevant. In these terms whether reality can be made conscious tends to be a contingent rather than a necessary matter, though the limits of possible consciousness may mean that some forms of reality are necessarily beyond what can be made conscious.

IV

The issue of consciousness can and should be de-idealized much further. Not only may the character of consciousness and how mind and matter presumably interact, if indeed they do, well be different in different species and periods of history and individuals differ according to personality, accidents, and surgery, but comparison of unusual cases with each other and with more common ones have helped allow experimentalists to gain vast insight beyond what mere speculation and conjecture seem able to contrive. Inded, imaginable real cases were they to occur could revolutionize the future. For example, could Siamese twins joined in the head survive and be capable of being studies and were they to share a common lobe, the possible discoveries which could result might stagger the most staid and cautious scientist.

Comparing common states of consciousness may also become very informative. For example, sleep, drowsiness, lack of concentration, shock, anger, depression, and other modes and moods could not only affect interaction with the physical world differently but could suggest that consciousness is capable of greater and lesser intensity as well as greater and lesser scope. (Virtually all analogies are inadequate. If consciousness were a "stream", it should be capable of being speeded up or reversed, and if it were a spotlight (perhaps the best analogy) the beam should be detectable. If the object of consciousness were a holographic picture[10] in the brain or a "ghost in the machine"[11] where is it located, why is it noticable, and what is the mechanism for turning it on and off?)

A man in shock may speak or move, control some voluntary muscles, but has trouble initiating plans or attending carefully to the world around him. Is he really conscious, and if so, is it a different form of consciousness (even apart from the issue of what he will subsequently remember)?

Also, there are different forms of loss of consciousness which can affect what is normally regarded as voluntary behavior differently. People who faint normally do not lose control of their bowels,[12] but people in a "drunken stupor" often do, and dead people always. I have not seen evidence on people in comas, on drugs, or under anasthesia. Sleeping people normally continue to control many voluntary muscles, but under some circumstances can lose control.[13] In short, not only may there be different ways and degrees of losing consciousness, but each one may appear to affect mind-matter interaction differently. In fact, the character of consciousness itself may both be and decompose differently in different species, individuals, and when affected by different causal factors.

V

No suggestion is intended that rearranging philosophical definitions and categories is always as futile as shifting deckchairs on the Titanic. The more fundamental the philosophical categories tampered with, the greater the impact on how we understand or misunderstand the world. Philosophy can be awesomely influential for those people who know how to detect such matters, even the current Neo-Scholasticism with its emphasis on logic, language, and universals.[14] But *if* rearranging definitions and categories is to revolutionize our understanding *to best advantage,* then I think it should be based on a much deeper historical and experiment-grounded understanding of central issues such as the varieties of consciousness and losing consciousness and how they may affect the brain and physical behavior.

The issue is not what to *call* consciousness, whether to label it "mental" or "physical", but to admit its probable reality, the reality of what is trans-conscious, and to understand how consciousness has evolved in different species and inividuals. Once these problems have been better understood, then we can return to the matter of general classification and decide what activities may best be considered physical, what mental, and what perhaps neither. We might even slowly gain insight into how disparate or incommensurable forces in this huge universe could influence each other, or on a more modest scale within the less atrophied parts of our brain.

ENDNOTES

[1] Sometimes the inflated notion of laws which many philosophers have is traced back to John Stuart Mill. Another possible source is the first two "Hauptsätze" of thermodynamics which are commonly mistranslated as laws in English. For most scientists a necessary law describes necessary relations between idealized variables, a statistical law statistical relations, and an empirical law empirical relations. The problem with the latter two types of laws is that it can be very difficult to draw valid deductions about particulars from them even under idealized conditions. Laws without variables can perhaps best be left to philosophers. Some constants may be non-idealized, but they are not laws.

[2] Galileo's Law of Falling Bodies, his time-distance law.

[3] Robert Boyle's pressure-volume law (also called Mariotte's law).

[4] Jacques Charles' volume-temperature law (also called Gay-Lussac's Law).

[5] A de-idealized approach is not merely needed when there are too many important variables to use laws, but whenever particulars would be understood not merely as examples of types (which by neglecting differences tends to idealize them) but as unique particulars.

[6] Darwin, Erasmus (1731–1802).

[7] Schischkoff, Georgi, (ed.), *Philosophisches Wörterbuch*, 21st edition (Stuttgart 1982), p. 261.

[8] Stegmüller, Wolfgang, *Hauptströmungen der Gegenwartsphilosophie*, Vol. I, sixth edition, (Stuttgart, 1978), p. 53 and Hans Joachim Störig, *Weltgeschichte der Philosophie* (Stuttgart $_{c}$1984), p. 622.

[9] Boring, Edwin G., *A History of Experimental Psychology* (New York 1957), p. 404.

[10] Pribram, Karl, *Consciousness and the Brain* (New York 1976).

[11] Ryle, Gilbert, *The Concept of Mind* (New York 1949), chapter 1.

[12] This may also depend on how long one is unconscious. If smelling salt cannot revive one, then presumably the condition is more serious than fainting. A pilot without oxygen is often said to "faint", but there is probably a more accurate word. Most people who faint presumably continue to breathe. Most people in a coma also continue to breathe. It is possible that everyday terminology cannot properly distinguish between control and loss of control over voluntary muscles. Nor is definition of the latter necessarily without problems, especially concerning breathing. Habitual and conditioned behavior may be voluntary without being volitional, which is different from non-voluntary or "instinctive" behavior which presumably can never be made volitional. Whether all habitual and conditioned behavior can be reversed and made volitional again is a moot point. Perhaps breathing contains both voluntary and involuntary aspects, with only some of the voluntary aspects being actively volitional even when we try to alter how we breathe.

[13] Strangely enough, some dreams can cause us to lose control. Perhaps most people have been embarrassed by such experiences at one time or another.

[14] Any approach which regards logic, language, or meaning as objective may be regarded as scholastic or artificial. If experience has taught us anything, it is that no two people use language or understand theories in exactly the same way. Not only do we all have different assumptions and patterns of usage as well as different vocabularies, but it is always possible to misunderstand logic, language, and meaning. It may sometimes be convenient to treat logic, language, and meaning *as if* they were objective (or consisted of universals), but whenever strict accuracy is important an historical, de-idealized approach is needed, which *distinguishes* between how different people use logic, language, and meaning.

* * *

THE POSITION OF THE THEORY OF MIND IN THE CATEGORIAL INTERPRETATION OF DIALECTICS

Izabella Nowak and Leszek Nowak
University of Poznań

1. In the tradition of Marxian epistemology there are two formulas which in completely different ways define the essence of human cognition:

(p) cognition is a derivative of social practice, and

(o) cognition is a reflection of reality.

Various lines of interpretation in that tradition in different ways refer to both those formulas. There is a complementary line which embraces them on equal terms: both those characteristics of cognition are to be equally indispensable and important, therefore none of them should be left out while building epistemology. There is also a stratification line which admits both those formulas, but does not accord to them equal significance: as the basic one, the formula about the subordinate character of cognition towards practice (praxistic position), or the formula of reflection (ontological position) are advanced. So both these orientations in the framework of the stratification line emphasize something different as the main characteristic of cognition—the element considered as primary by the praxist, is treated as secondary by the ontologist, and vice-versa. This has such a methodological sense that while building epistemology the second is left out in the first approximation. The follower of the priority of the (p) formula therefore promises himself to take into account the connection between cognitive and ontological structures, after he has finished with the relation: cognition-practice; whereas the supporter of the priority of the (o) formula will first characterize the relationships of cognitive structures and the principles of their development to the ontological structures and the principles of their movement, in order to supplement and modify the obtained results by taking into account the elements of a pragmatic nature. This also has such a systematic sense that for the adherent of the praxistic formulation, what is assumed by Marxist epistemology is the theory of historical materialism, whereas the adherent of the ontological position assumes objective dialectics. The role of dialectics in the framework of praxistic epistemology, as well as the role of historical materialism in the framework of ontological epistemology becomes secondary, and usually—simply vague (e.g., when the praxistic approach to epistemology is accompanied by the praxistic interpretation of dialectics itself). Thus sometimes a nihilistic line comes into being in the framework of Marxism, which eliminates the (p) idea—variant anti-praxistic—or the (o) idea—anti-ontological variant. This occurs against the very tradition which after all evidently embraces both formulas, and—which is more important—against reality, because in fact cognition reflects reality, and simultaneously depends on social practice.

In the present article we would like to make a distinction between the problems of the theory of mind and the problems of the theory of cognition, show the position of the theory of mind in the framework of the categorial interpretation of dialectics, and also prove that this approach does not evoke nihilistic (or else liquidating) tendencies towards the (p) idea, which positively distinguishes it among some alternative interpretations of Marxist epistemology, which are discussed at the end of this article.

2. Dialectics, according to the categorial interpretations, is distinguished from among other essentialistic doctrines in that it is a dynamic essentialism. Essentialism asserts that there is a hierarchy of factors affecting each factor and the following factors are distinguished: primary and secondary. For each factor, therefore, there is a regularity relating it to the element(s), primary for the factor. That regularity is manifested in certain forms which express a combined effect on the given factor both of the primary as well as secondary magnitudes. What

distinguishes the categorially interpreted dialectics from among other essentialistic doctrines is the thesis that the essential structures of phenomena (factors) are changing. One should differentiate between the transformations of the essential structure of a phenomenon (qualitative changes) and the developments of that structure (quantitative changes). The former consists in that the set of primary factors is changing, whereas the latter in that the hierarchy of secondary factors undergoes changes with identical primary factors. If a qualitative change, therefore, takes place, then the very phenomenon (i.e., the fact consisting in that a given factor appears in a definite intensity on some object) is subject in comparable moments to different regularities. If, however, a quantitative change occurs, then it is subject to the same regularity, but that change is manifested differently in comparable moments, its forms are different.

The categorial history of a phenomenon is the series of its essential (quantitative or qualitative) changes. The historical process, therefore, in the categorial understanding consists in changes of the essential structure (developments or transformations). At each stage of the categorial history, a phenomenon is subject to various dependencies (its nomological structure is different)—or else to different regularities (in case of qualitative changes), or to different forms of the same regularity (in case of quantitative changes).

The principles of objective dialectics are interpreted as certain general theses on the form of the categorial histories of phenomena. Consequently, the principle which says about transformation of quantity into quality is understood as a thesis that the essential structure of each phenomenon is subject first to a series of developments, and only then to transformation; whereas the principle of double negation is interpreted as a statement which proclaims that there are not such transformations of essential structures, where a new structure does not have certain factors in common with the former one, therefore some range of continuation still remains.

3. The above outlined ideas of the categorial interpretation of objective dialectics are not precise enough[1] and incomplete (since they outline only the most idealized model of that conception).[2] However, they permit catching the moment—important for the present article, viz. the adequacy, correspondence of cognitive and ontic structures. The structure of the idealizational theory of a phenomenon is in some way a reflection of the essential structure of that phenomenon. The hierarchy: regularity and its subsequent forms corresponds to the hierarchy: the idealizational law and its subsequent concretizations. If, therefore, it is true that the mehod of idealization is the key method of constructing scientific theories, then it may be said that the structure of a theory is not an arbitrary creation of mind, but it reflects the internal structure of the phenomena explained by the theory.

Similarly, each idealizational theory passes through a number of developments—ever more numerous sets of idealizing assumptions are developed where, however, the same basic laws are retained, and then those laws are subject to concretization, among others with regard to new assumptions. A former law becomes more abstract in the framework of a new, more developed theory, but at the same time better concretized. After a number of developments, however, the theory is subject to transformation—a basically different picture of the essential structure with a different set of primary factors is assumed, a new idealizational law is established and gradually concretized. Never are all factors abandoned formerly treated as essential—some of them are still retained consequently, some elements of the former theory (discovered in its constituent dependencies) are incorporated into the new theoretical structure. It is apparent therefore, that as the history of a phenomenon is composed of a series of quantitative changes leading to qualitative ones, so the history of a scientific theory is composed of a series of developments of the initial theory, leading to its transformation. If the thesis that theories have idealizational structure is an equivalent of the thesis of essentialism, then the thesis that they undergo changes by developments and transformations is an equivalent of the principle that quantity transforms into quality; and the thesis that none of the new theories completely breaks away from the former, but constitutes at the most its transformation (i.e., retains some former dependencies at least in the form of concretization of new laws), is the counterpart of the double negation principle.

4. Even with the foregoing considerations[3] one may notice correspondence between theses of dialectics interpreted categorially and epistemological theses concerning the structure and development of scientific theories. Both the internal structure of a theory as well as the principles of its movement are isomorphic in relation to the structure of phenomena and principles of their movement.

Naturally that may only occur for very strong idealizing assumptions. We will call such a cognition as sovereign which is capable of discovering the idealizational theory for each value F, including the absolute truth, and discover immediately without correcting earlier suggestions; a mind capable of such cognition we will call perfect. (It need not be said there are no perfect minds, as there are no perfect Chomskian speakers−recipients, whereas examination of such constructs may be useful; this cannot be decided a priori: such examinations should simply be carried out). And so for the sovereign cognition among others the following principles of subject dialectics are true:[4]

(1) each local history of any thought history includes transformations preceded by a set of developments,
(2) each local history of each thought history includes exclusively transformations of the first type.

Those principles are the counterparts of the ontological principles saying that quantity is transformed into quality and the principle of double negation from objective dialectics, and so are other theses of subjective dialectics. Consequently, it may be stated that subjective dialectics needs the following condition which we will call a categorial interpretation of the formula of reflection:

(o') for each F the structure of the mental category of F type is isomorphic to the structure of the subjective category of F type, and the thought history of the F factor is isomorphic to the objective history of that factor.

Naturally, the conception of sovereign cognition is an idealization of the conception of human cognition. In real cognition we deal with the process of approximating the reality (never all concretizations are carried out effectively, a number of concretizations always ends with approximation), therefore the subsequent theoretical proposals do not contain theses which would claim the role of absolute truth, at the most they formulate their approximate equivalents more and more precisely with the progress of science. In real cognition, original theoretical proposals are always corrected (new secondary factors are introduced not because they were not acting before and now come into being, but simply because they were not recognized before), consequently each theory initiates a series of corrections. And so one assumes in an idealizing way about the sovereign cognition, and correspondingly, about the perfect mind that it is capable of formulating and not only approximating the absolute truth about a given phenomenon, and that it is carried out without correcting itself (consequently that self-correcting series consists of one element in this instance).

The theory of mind (or the theory of sovereign cognition), therefore, will be that part of epistemology which is binding with the idealizing assumption of the sovereignty of cognition. Before we will try to define the position occupied by a so-conceived theory of mind in categorial epistemology, there will be some comments on its constituents.

5. The theory of mind embraces first of all that model of cognition which is defined in the categorial interpretation of dialectics as subjective dialectics[5] with the principle of idealization and also the principles of subjective dialectics. Subjective dialectics is based not only on idealizing assumptions about the sovereignty of cognition, but also on all idealizing assumptions of objective dialectics. Consequently, if objective dialectics[6] is based on the assumptions idealizing the structure and development of the external reality, then subjective dialectics must make assumptions which constitute that ideal model of reality. And when those assumptions are gradually lifted in the subsequent models of the categorial ontology, and the ideal type of the external reality becomes ever more realistic, then the picture of the world examined by epistemology is thereby enriched, and as a result the theses of subjective dialectics are corrected and supplemented. And so, e.g.−referring to the already performed concretization of

objective dialectics, one should state (see point 2) that it is binding for, among others, two idealizing conditions. One of them says that material objects are subject exclusively to the process of essential differentiation (i.e. objects belonging to one categorial class, may belong in the future to one or more categorial classes)[7], and the other that factors are mutually essential at most in a one-directional way (i.e., if one is essential for the other, that other is not essential for the first one).

Consequently, removing the first of those conditions a unification of objective categories is introduced, as a reverse operation to their dissociating; a differentiation is made between historical and physical time; the conceptual range of subjective history is widened (distinguishing among others the history with unifications, history aggregatively differentiating, aggregatively unifying, etc.); formulations of the principles of objective dialectics are corrected and a new ontological thesis is added to those principles—the thesis about unevenness of movement: "each objective history is a history aggregatively differentiating and simultaneously uneven in terms of time".[8] That way the picture of cognition examined in epistemology must undergo adequate changes—nevertheless, cognition is related now to the ever more complex ideal type of external reality. Since the real processes cover not only the essential differentiating but also unifications, consequently the sovereign subject—and the assumption about sovereignty is constantly in force—must reflect those processes in his constructs. A unification of mental categories, therefore, is introduced,[9] the conception of the history of thought is appropiately widened etc., first and foremost the principle of the concreteness of truth is introduced: if cognition is sovereign, then consequently:

(3) the thought history of each factor is aggregatively differentiating, and differentiated in terms of time.

Since the time when the phenomena develop unevenly in different ranges of their existence, thereafter the sovereign cognition must designate for those different ranges different idealizational theories for the same values; in this sense "the truth is always concrete". But the thesis (3) does not belong to subjective dialectics, since it is not an epistemological equivalent of the principles of objective dialectics, it is instead an equivalent of the thesis about unevenness of movement of one of the derivative models of categorial ontology.

Subsequently, lifting the second of the mentioned idealizing conditions, the conceptions of coupling and system (a set of mutually coupled factors) are introduced, which enables reconstructing the thesis about an all-embracing union, saying that material reality is a system. And again: after the ideal type of reality is examined in the present, a more realistic model of the categorial ontology is to be reflected—and the idealizing assumption about the sovereignty of cognition is still in force—then the cognitive constructs should be appropriately enlarged. Consequently, one introduces appropriately the couplings between the idealizational theories, and conceptions of the theoretical system treated as a set of such theories of which every two are coupled. This enables reconstructing the procedures of analysis and synthesis in the specific sense of those terms, included in the Marxist tradition, and first of all—reconstructing the thesis about the system character of truth: if the cognition is sovereign, then

(4) the absolute truth about a given factor cannot be established without establishing such truth about every other factor.

The thesis about the system character of truth is an epistemological equivalent of the ontological thesis about the universal nexus.[11]

And so theses (1) and (2) are examples of theses of subjective dialectics and consequently of the theses from the most idealized models of the theory of mind. Whereas the theses (3) and (4) constitute examples of the more realistic epistemological models, but still belonging to the theory of mind since they all meet the categorial condition of reflection (*o'*). Those models of epistemology, beginning with subjective dialectics, which fulfil the categorial interpretation (*o'*) of the formula: "cognition is a reflection of reality", constitute the theory of mind. Consequently, the theory of mind comprises the concretizational sequence of those epistemological models which retain isomorphism of ontological and cognitive structures.
Such a concretization of the epistemological model, which goes beyond that isomorphism does

not belong to the theory of mind. It may easily be established which type of concretization goes beyond that isomorphism, and therefore satisfies the condition of reflection (*o'*)—namely such which refers to the factors specially related to the character of the cognitive means employed by people, structure, of cognizing societies, etc. In short, beyond the theory of mind there proceeds such a concretization of the models of that theory, which assumes lifting the idealizing condition on the sovereignty of cognition. For that condition is in fact a decision not to take into consideration the cognitive limitations imposed on us 'people' by the structure of the perception apparatus, the nature of notions that we use, the methods for passing information, limitations—but also of amplifications—of the process of cognition, resulting from its numerous involvments in social practice, etc. Whereas beyond the cognition of reflection (*o'*) these concretizations are not brought out which result from ontological simplifications, consequently from the idealization of the object of research.

That way, in the framework of epistemology one may distinguish the theory of mind—a sequence of models beginning with subjective dialectics, which while presenting an ever more realistic (as concerned with the ever more realistically presented reality) pictures of cognition, retain the condition of reflection (*o'*), and the theory of human cognition—a sequence of models, based on an ever broader taking into account of circumstances specific for our cognition, and therefore on the ever further weakening of the condition of reflection.

That structure of epistemology is presented graphically by Fig. 1.

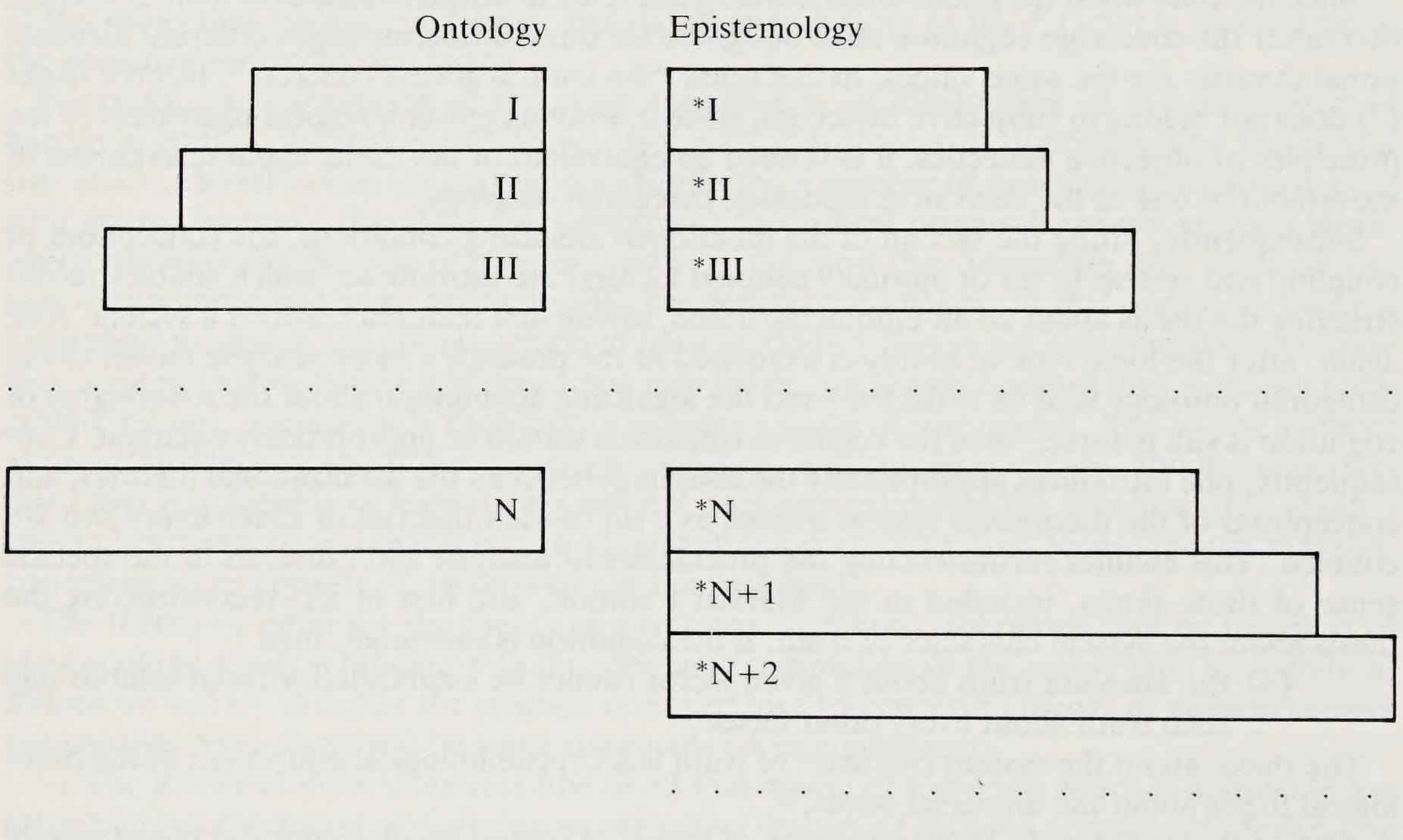


Fig. 1 Categorial ontology consists of the models I, II, III . . ., N, where the model I is objective dialectics. The categorial epistemology consists of the models *I, *II, *III . . ., *N + 1, *N + 2, *N + 3 . . ., where model *I is subjective dialectics. The categorial epistemology is composed of the theory of mind (the first epistemological models retain the condition of reflection (*o'*), and the theory of human cognition (further models).

6. The differentiation between the theory of mind and theory of human cognition leads to some consequences worth, as it may be assumed, noting. The first is the possibility of reinterpreting the known distinction between normative and descriptive epistemological considerations. The models of the cognitive process from the theory of mind present a certain pattern of cognition—as far as the structure of cognitive constructs and the principles of their development are concerned. Whereas the models of the cognitive process from the theory of human cognition present more realistic approaches which are simultaneously more distant from the pattern of sovereign cognition. Consequently, it could be said, that the theory of mind is "normative", and the theory of human cognition—"descriptive", but only when basing this differentiation on completely different principles. For neither the theory of mind formulates "norms", nor the theory of human cognition formulates "sentences". They both formulate some sentences, viz.—the idealizational statements. Whereas formulating norms in the theory of mind would be simply an irrational activity in terms of norms creation, for it is known that implementing the patterns of sovereign cognition is impossible; nobody establishes norms demanding states of affairs which are impossible to be carried out.[12] On the other hand, establishing idealizational statements which are to define the structure and development of sovereign cognition (i.e. fulfilling the condition of reflection (*o'*)) permits defining our cognition in order to represent the nature and movement of reality itself. It plays, therefore, the role which is similar to the one played by, e.g., the structure of the system of the omniscient assertion logic, where the following formula is an axiom: "*p* is asserted then and only then, when *p*"; the omniscient assertor recognizes all true sentences and only true sentences—consequently knows everything and does not go astray.[13] The system of omniscient assertion indicates a certain pattern of perfect assertion in both senses of the word—idealized and correct. The theory of mind aspires towards something similar.

7. The theory of human cognition includes also, among others, models presenting the influence of social factors which weaken the sovereignty of cognition. This refers both to the internal factors related to the functioning of cognizing societies, as well as external refering to the role played by the products of those societies in broader social contexts. That type of model of sociology, deals with completely different issues than the models of the theory of mind, dealing not so much with correct cognition as with the conditions for attaining correct cognition by real cognition. What true cognition consists in, what are the conditions for transforming a theory into a subsequent one with retaining at least the same degree of truth, etc. —these are the issues undertaken in the framework of the theory of mind. Why physics seems to attain today a higher explanatory power than, let us say, economy, why the conditions of correspondence of theory are abused in economy, etc.—these are issues undertaken in the sociological theory of cognition. What is the structure of a scientific theory (whether it is defined by the relation of consequence, or concretization, or any other)—it is an issue belonging to the theory of mind. As to which social theories express the perspective of the class of owners, and which of the classes of rulers or priests[14]—is an issue belonging to the sociological theory of cognition. And so on.

Consequently, one should distinguish issues related to the characteristics of cognitive patterns—issues of the theory of mind; and issues related to the attainability and the social role of those patterns—functional issues belonging to the theory of cognition. Primary is that the issues of the first kind are in a quite definite sense pre-eminent in relation to the functional issues: one must assume solving the former in order to raise the latter at all. One must assume some conception of theory in order to be able at all to ask the question which scientific theories express the interests of which classes of disposers of material means. This corresponds to the order of the theory of mind and theory of human cognition (among others, therefore, to the sociological models of that theory) in the framework of epistemology. The mentioned order is an expression of that in the categorial interpretation of dialectics the formula of reflection (*o*) is pre-eminent in relation to the formula of subordination of cognition to practice (*p*). One must first know what cognition is, before one asks how it is affected by social practice and vice versa.

That way it is more apparent why in the framework of the categorial interpretation of dialectics—from among the mentioned in point 1—the stratification line in the ontological variant is assumed. The very formula of subordination (p) is an answer to the functional question. On the other hand, the formula of reflection (o) is an answer to the question belonging to the theory of mind.

ENDNOTES

[1] For more on the subject L. Nowak, *U podstaw dialektyki Marksowskiej. Próba interpretacji kategorialnej (At the Basis of the Marxian Dialectics. An Attempt at a Categorial Interpretation)*, (Warszawa 1977), and "The Structure of Marxist Dialectics. Towards a Categorial Interpretation" *Erkenntnis*, vol. 11, No. 3 (1977). See also papers by I. Nowakowa, S. Magala and L. Nowak in *Categorial Interpretation of Dialectics, Poznań Studies in the Philosophy of the Sciences and Humanities*, vol. 2, No. 4 (1976), Amsterdam.

[2] For further models see: L. Nowak, *Wykłady z filozofii marksistowskiej (Lectures on the Marxist Philosophy)*, vol. II, *Ontologia i Epistemologia* (Poznań 1978).

[3] For more on the subject see: I. Nowak, *Dialektyczna korespondencja a rozwój nauki (Dialectic Correspondence and the Development of Science)*, (Warszawa-Poznań 1975), and *Z problematyki teorii prawdy w filozofii marksistowskiej (Problems of the Theory of Truth in the Marxist Philosophy)*, (Poznań 1977); L. Nowak, *The Structure of Idealization. Towards a Systematic Interpretation of the Marxian Idea of Science.* (Dordrecht 1980).

[4] See: L. Nowak, *U podstaw. . .*, p. 250 and L. Nowak, *The Structure of Idealization.*

[5] Nowak, L., *The Structure of Idealization.*

[6] See works quoted in endnote 1.

[7] Ławniczak, J., "Rozszczepienia i unifikacje. Próba rozwiniecia kategorialnej interpretacji dialektyki" ("Splittings and Unifications. An Attempt at Developing a Categorial Interpretation of Dialectics"), in Założenia dialektyki (The Assumptions of Dialectics), (Warszawa-Poznań 1977); L. Nowak, *Wykłady. . .*

[8] Nowak, L., *Wykłady . . .*, p. 21.

[9] Ibid., p. 86 and the following.

[10] Ibid. and I. Nowak "The Principle of Universal Nexus in the Categorial Interpretation of Dialectics", in *Categorial Interpretation of Dialectics, Poznań Studies in the Philosophy of the Sciences and the Humanities,* vol. 2, No. 4.

[11] Nowak, I., Nowak, L., "Zasada wszechzwiazku i systemowość prawdy" ("The Principle of Universal Nexus and the Systematic Character of Truth"), in *Odkrycie, abstrakcja, prawda, empiria, historia a idealizacje (Discovery, Abstraction, Truth, Empiricism, History and Idealizations)*, ed. A. Klawiter, L. Nowak (Warszawa-Poznań 1979).

[12] Nowak, L., Próba metodologicznej charakterystyki prawoznawstwa (An Attempt at Characterizing Jurisprudence), (Poznań 1968).

[13] Marciszewski, W., Podstawy logicznej teorii przekonań (Foundations of the Logical Theory of Convictions), (Warszawa 1971).

[14] See L. Nowak, *Property and Power. Towards a non-Marxian Historical Materialism* (Dordrecht 1983).

* * *

THEIR SKULLS AND OUR THOUGHTS

William James Earle
Baruch College Cuny

These remarks were occasioned by Rodney Needham's "Skulls and Causality." According to Needham, anthropologists have interpolated a link between the taking of heads and the supposed benefits of that practice; they have written as if their informants in the field had answered the question 'But *how* does the taking of heads procure benefits?' whereas the informants appear to find this question puzzling or unintelligible. I assume that Needham is right about what the ethnography will not sustain; based on the ethnography we should say only "The taking of heads secured well-being, and that was that".[1] One can sympathize with Needham's effort not to impose upon the natives our explanatory standards and categories; we have no reason, *a priori,* to think they share our beliefs about causality and causal explanation. But what do they believe? I am not, of course, going to try to answer this question (that belongs to anthropology), but to consider one purely philosophical, or conceptual, problem that arises in the attempt to ascribe beliefs. There may be other problems too, but one is enough to discuss in this brief compass.

The problem to be discussed arises in the following way. Suppose we set out to describe, without the imposition of anything alien, the native *Lebensform*, that is, we begin in the spirit of Needham which is also the spirit of Wittgenstein as evinced in the following characteristic remark: "Nur *beschreiben* kann man hier und sagen: so ist das menschliche Leben".[2] We can indeed describe a great deal: exactly how the natives gather the skulls of their dead enemies, how and where they preserve them, and what they, from time to time, do with them. We are being like, or trying to be like, a camera; and of course we might actually film the head-hunters. The film might be the most neutral, least deforming, way of showing what their life is like. But *showing* isn't saying: the perplexity we should feel standing at the edge of the clearing, we shall feel viewing the film if it really is a neutral conveyer of the scene. Whether we are on the scene or watching the film, we want to know what is going on, what the natives think they're up to, what they believe. The flattest, most external, least theory-beholden description of the natives (a kind of Hemingway or *nouvelle roman* flatness) will be the verbal analogue of the film: precisely not something that can be rendered false by anything the natives turn out to believe, but (like the film itself) it won't touch our perplexity.

Why can't we, to deal with perplexity, extend the descriptivist project to encompass native beliefs? This is where the conceptual problem I referred to above emerges. An external description of an agent's action [so we may define it] is a description that cannot be rendered false by anything the agent turns out to believe; an external description of a person's belief [analogously defined] is a description that cannot be rendered false by anything the person turns out to believe. (N.B., 'false', here, means 'false of' or 'inapplicable'.) It is sometimes claimed that external description even of action is impossible, though it is not clear why we can't frame such descriptions, 'thin' and uninteresting as they may inevitably be judged. Given the definition of 'external description', there can be no similar uncertainty about whether external description of belief is ruled out: external description of belief is simply logically impossible.

This has (to be clear) nothing to do with the question of how good we can become at ascribing beliefs from the external point of view, that is, how good we can become at ascribing beliefs based only on external descriptions. 'The woman rises from her chair, goes to the window and looks out; then she opens a closet, takes out a raincoat which she puts on and buttons up, walks through the hall where she picks up an umbrella, and leaves the house'. I would

count this as an external description; no odd beliefs she has about what she is up to will make this description incorrect. (It can, of course, be incorrect for other reasons: the 'raincoat' may be a bulletproof coat, the 'umbrella' a gun, her standing before the window a signal to the other terrorists.) With this external description and our knowledge that it *is* raining, we say that she believes that it is raining; if we know it is not raining, we say that she believes that it is going to rain. In either case we have probably made a successful ascription of belief. Nevertheless our ascription of belief can turn out to be mistaken because of what she actually believes ('It's not going to rain, I'm not water-soluble anyway, but I might as well seem Miss Prudent to the executors of my trust who may spring for more money if they think I'm a sensible sister, etc.') whereas the description it was based upon will stand.

We ought (for the sake of further clarity) to add an utterance to the little scene we described. This time the woman, as she turns from the window, utters 'It is raining' (perhaps there is a maid, or a terrorist, standing in the part of the room we cannot see). To describe an utterance is to quote it. A quotation is like a taperecording. To describe the audio addition to our little scene, we simply say 'and then she said "It is raining"' or we say 'and then she said...' and press the button of our machine. Since quotation keeps us external, we are safe in saying she said 'It is raining', that is, uttered those words, no matter what she thinks she did or was up to. Ordinarily, the transition from direct quotation (she said 'It is raining') to *oratio obliqua* (she said that it was raining) is as unremarked as it is unremarkable: a minor shift in style of reportage.

There are circumstances, however, in which we cannot move beyond quotation. Let us suppose that the head-hunting natives interviewed by Needham and other ethnographers, when pressed for an explanation of *how* the taking of heads works, will only respond 'It just works'. We can of course say, not only 'They say "It just works"', but 'They say that it just works' and we can even go on to say 'The natives believe that it just works'. The move beyond quotation in the two sentences last mentioned is, I want to argue, merely apparent. In order to see this we have to ask ourselves how we understand 'X says that p' and how we understand 'X believes that p'. I say 'how we understand S' rather than 'what S means' because what we are looking for is a mapping of sentences to understandings [or construals or something of the sort], not a sentence to sentence mapping which, however accurate, would not explain how we understand either one. I suggest that we understand 'X says that p' as *X says what we would say in uttering 'p'* and understand 'X believes that p' as *X believes the belief we would express in uttering 'p'*. In offering these construals, I am simply borrowing the kernel of a view recently elaborated by Stephen Stich. According to Stich, when I say 'X believes that p'"...I am performing a little skit, doing a bit of play acting. The words following the demonstrative 'that' in the sentence I utter constitute the script for the play." I am then saying [says Stich] that X"... is in a belief state content-identical to one which would play the central role in the causal history leading up to my play-acting assertion, were that assertion to have been made with a typical causal history."[3] What is at issue in Stich's account, as he puts it, is a "relation of content identity that links the belief state of the believer with the hypothetical or counterfactually characterized belief state of the attributor."[4] Stich's view contributes to my argument in the following way. Where 'p' is 'It just works', there simply is no relevant belief state of the attributor, no typical causal history, no possible performance of the skit of ascription. It is obvious that, on the account offered here, my understanding of either 'X says that p' or 'X believes that p' is dependent on my prior understanding of plain 'p'. Where I really draw a blank for 'p' (as I do for 'It just works'), I can certainly say 'X says that p' or 'X believes that p', but *the most I can mean* is 'X says "p"' or 'X believes something he expresses by uttering "p"'. In my own society I may—at least—occasionally suspect that X is as blank as I am, that he is quoting 'p' to himself without being aware of it, without realizing that he has failed to give 'p' a real use, and that in fact he has no belief. Some of these will be the cases of 'inadvertent nonsense' made infamous long ago by the logical positivists. In the context of anthropology, we normally suppose that there are beliefs. Access to those beliefs, we must—perhaps sadly—conclude cannot be achieved through mere good will, imaginative sympathy, or prejudice-shedding.

ENDNOTES

[1] Needham, Rodney, "Skulls and Causality", in: *Against the Tranquility of Axioms* (Berkeley, 1983), p. 79.
[2] Wittgenstein, Ludwig, *Bemerkungen über Frazers Golden Bough* (Nottinghamshire 1979), p. 3.
[3] Stich, Stephen P., *From Folk Psychology to Cognitive Science: The Case Against Belief* (Cambridge Mass. and London 1983), p. 84.
[4] *Ibid.*, pp. 84–85.

* * *

WEAK METHODS AND SEMANTIC RICHNESS

R. Bhaskar
IBM Thomas J. Watson Research Center, Yorktown Heights, New York

0. Introduction

Artificial intelligence seeks to discover invariants in intelligence and intelligent behavior. One of the chief elements of intelligence is knowledge. This paper is about invariants in the use of knowledge. One approach to this problem is to postulate an invariant and empirically examine specific instances of actual behavior. It is then possible to decide, formally or informally, if the example supports or reflects the claims of a particular invariant. One example of an invariant is a general problem-solving method. Usually, general methods have to be substantially modified before they can help in performing cognitive acts, such as problem solving, especially in domains of special knowledge. These modifications can be called, "semantic adaptations."

To say that a particular method in a particular domain is a "semantic adaptation" is to assert the following: (1) there is a general method available for use in many different domains; (2) the method used in the particular domain being studied is similar to the general method, but not identical to it; (3) whatever differences exist between the two methods, the general or the particular, are caused by the role of specific knowledge elements in the domain. An example from everyday life will illustrate this in a simple way.

If you want a theory of traffic patterns of automobiles in the United States, it is a good idea to begin by noting a few laws that prevail over the entire country: the speed limit is 55 mph, you drive on the right, pass on the left, etcetera. Besides this, some areas have their own rules: you see signs that say DETOUR, or CONSTRUCTION AHEAD, and so on; it is easy to conclude that these signs affect traffic patterns locally. It would be nice if these *national* rules along with a few local rules were to be sufficient for understanding traffic patterns; in practice however, there are also large cities that have their own rules. To study traffic patterns everywhere, it is necessary to understand that intra-city traffic is important and large in volume; it is also necessary to realize that traffic patterns in the city cannot be understood just by understanding the rules. Instead, understanding city traffic requires knowing more about the city itself: where people live, where they work, the direction and timing of the morning rush hour, and so on. Within the city, studying traffic patterns is a complex matter, requiring study of matters that are not related to traffic in an obvious direct way, such as the economic and social geography of the city; these matters may vary from city to city.

A scientist studying cognition is in much the same position as the traffic theorist. It is definitely a good idea to begin by assuming some generality; a few general processes, a more or less uniformly organized memory and a control structure that facilitates complex information processing. Sometimes local problem conditions may allow shortcuts that are obvious at the situation; these local conditions apart, the picture is pretty clear, or so one hopes.

In practice of course, the situation is a little less cheerful; there are cognitive equivalents of the cities in our example. We call these cities, *domains*, and in understanding the patterns of cognitive processes on the whole, these domains play an important role. Here again, our example helps us to understand. We cannot understand the traffic patterns of a city just by knowing traffic laws in the city; instead, we need to learn something of the economic and social geography of the particular city. We need to learn something about the downtown area; who goes there, when or how do they get there. Analogously, we learn about cognitive processes by understanding the concepts of the domain. To understand this, we shall explore the

domain with a general-purpose cognitive vehicle, the weak method (Newell, 1975). In this paper, we shall see how a domain looks when a weak method is used to solve problems in it.

Solving problems in a semantically rich domain with a weak method is a lot like driving in a strange city equipped with a map, but no special knowlege of the city. It can be done; the way it is done is probably not the simplest or easiest way. The second time around, even a novice may change to a route that does not necessarily seem, on the map, to be optimal; perhaps it does not even appear on the map.

Safe driving requires knowing your car; a weak method is a fairly general problem-solving method that is quite flexible about its inputs. A fairly common weak method, of which several brands exist, is the one generically termed "working backward." The rationale of this method is quite simple: analyze where you are, where you want to be, and plot (plan) a series of steps backward from your goal to the current (sad) state; if the plan is sufficiently detailed, and complete, actually executing the plan can be made quite trivial. However, things can go wrong. There may be error in calculation, in information, or perhaps in judgement. The weak method is now useful simply because it is general and can be used for error-recovery procedures.

1. Example 1

One place where working backward can be especially helpful is in solving a system of equations one by one. Equations of this sort appear in engineering thermodynamics. In a first course on thermodynamics, many problems involve the following four equations:

$$\dot{Q} - \dot{W} = \dot{m}\left(\frac{1}{2g_c}(v_2^2 - v_1^2) + C_p\,(T_2 - T_1)\right) \tag{1}$$

$$\varrho_1 A_1 v_1 = \varrho_2 A_2 v_2 \tag{2}$$

$$P_1 = \varrho_1 R T_1 \tag{3}$$

$$P_2 = \varrho_2 R T_2 \tag{4}$$

where

$\dot{Q} \equiv$ rate of heat input

$\dot{W} \equiv$ rate of shaft work output

$T \equiv$ temperature

$A \equiv$ area

$P \equiv$ pressure

$\varrho \equiv$ density

$v \equiv$ velocity

There are twelve variables and four equations. Since the equations are independent, it is possible to solve all four equations if eight variables are given. Consider for example, the following problem:

The working fluid of a flow system is air. The work done 592 lbf.ft/second. The inlet temperature is 46.1 degrees Fahrenheit. The outlet velocity is 15.8 feet per second. The heat input is 154.97 BTU/sec. The inlet area is 4.6 square feet. The inlet pressure is 16.8 psia. The outlet area is 36 sq.ft. The outlet specific volume is 0.3 cubic feet per lbm. What is the outlet pressure?

Ordinary application of the working backward method results in the equations being solved in the following order:

first Equation (3) for ϱ_1
then, Equation (2) for v_1
next, Equation (1) for T_2
finally, Equation (4) for P_2

In point of fact, a human who was asked to solve this problem did not solve the equation in this order. Instead he first wrote down Equation 1 (incidentally, equation 1 is the first law of thermodynamics), then selected and formulated the equation of conservation of mass, and *then* as a *third* step, wrote down and solved Equation (3), the equation of state at inlet. This step is third, rather than the first, as dictated by the standard method of working backward. This sequence is of course different from that predicted by a simple application of the process of working backward. This particular variation is repetitive, in the sense that the I law is consistently used as the basic equation that drives the behavior during the entire problem solving episode. The actual behavior of the human has been discussed in detail by Bhaskar and Simon (1977), especially pp. 209–212.

This particular sequence is actually a simple variant of the method of working backward. It is produced by adopting the following rule: *regardless of the problem, solve the energy equation first; after that, solve whatever equation(s) are necessary to calculate the particular variable the problem requires.* Solving the energy equation is, of course, another example of equation-solving and is itself accomplished by working backward, this time from the energy equation. Thus the "variant" we speak of essentially consists of breaking up the problem required into two separate problems, each of which is solved by a direct application of the method of working backward.

The variation from a "canonical" backward approach is not a random variation; rather it is specifically caused by the importance of the first law of thermodynamics in the semantic topography of the domain. This variation is not the only possible one; the equations of state (3, 4) are the expression of an important physical law (Gibbs' rule) as is the conservation of mass (equation 2). The central role of equation (1) cannot be predicted from theory; instead it can only be discovered by careful empirical exploration of cognition in the domain. Furthermore, it is quite important to know the equation-solving sequence for a theory of cognition that is satisfactorily complete.

2. *Example 2*

Physics provides another example. In elementary physics, the following equations are taught:

$$v_t = v_0 + at \tag{1}$$

$$s = \bar{v}t \tag{2}$$

$$v_t^2 = v_0^2 + 2as \tag{3}$$

$$s = v_0 t + 1/2at^2 \tag{4}$$

where

$v_0 \equiv$ initial velocity

$v_t \equiv$ velocity at time t

$\bar{v} \equiv$ average velocity

$s \equiv$ distance traversed in time t

$a \equiv$ constant acceleration

For many problems, one of these equations is sufficient. They offer a systematic and "guaranteed" path to the solution. The expert, however, does not appear to use these equations, but instead a simpler form of the following equations using average velocity as a variable. Consider, for example the following three problems:

1. How long will it take a freely falling body to acquire a speed of 160 feet per second and *how far will it have fallen during that time* (Problem 6, Simon and Simon, 1978)?

2. A car moving on a straight road increases its speed at a uniform rate from 20 feet per second to 30 feet per second in 20 seconds. *How far did it move* (problem 7, Simon and Simon, 1978)?

3. A bullet shot vertically upward has an initial speed of 1600 feet per second. How long did it take before the bullet stops rising? *How high does the bullet go during this time*? (Problem 18, Simon and Simon, 1978).

In each of these, a certain distance is required to be calculated; two equations are available for this: equations (2) and (4), above. The expert consistently uses equation (2) and the novice, equation (4). The behavior of the expert here is also a special case of working backward. In this case the variant is produced by the rule: *calculate the average velocity first, then solve for the variable that is required by the problem statement*. This again, is a special case of solving the original problem by solving two separate problems by working backward. Thus it is a variant of working backward. Here the first subproblem has its goal the calculation of a particular *variable*. This variable is important because it allows physical intuition to play a role in solving physics problems (Simon and Simon, 1978).

As before, the point is that this choice of equations is not accidental. Where equation (4) is derived from principles of classroom physics, equation (2) can possibly be claimed as physical intuition, and indeed is so described by Simon and Simon (1978).

3. Conclusion

In two domains, experts use variants of the method of working backward; in one, thermodynamics, the variant consists of solving a particular equation, the *energy equation*, and then proceeding with the textbook problem; in the other, physics, the variant consists of solving for a particular variable, *average velocity*. Both of these are specific modifications of the weak method called "working backward." These modifications arise because of the the semantic topography of the respective domains. Understanding the laws of semantic topography is an important research question.

The equation-solving sequence is not a superficial description of the behavior of an expert; it reflects the effects of many different components, such as memory organization, problem

solving, planning and goal formation processes, problem representation, and so on. For many practical purposes, especially those connected with learning and instruction, understanding this behavior may be the only way in which a change in behavior can be readily observed. That is to say, the equation-solving sequence appears to be systematically different for experts as compared to novices; the equation-solving sequence used by a particular human subject can be readily observed; thus it is easy to see if something has been learned or taught. To do this, however, the expert's equation-solving sequence must be recognized and understood.

What we are calling the equation-solving sequence is, of course, just one instance of an element of behavior that can become important in understanding cognition. In other domains, other elements will appear. What these elements will be cannot be determined by theory, but only by exploring the way the knowledge in the domain has accrued and has come to be used. This can only be done by careful empirical studies of cognition in specialized domains.

ENDNOTES

I am grateful to J. R. Hayes for comments on an earlier draft.

Ayer, A. J., *Logical Positivism* (New York 1959).

Bhaskar, R. and Herbert A. Simon, "Problem Solving in Semantically Rich Domains: An Example from Engineering Thermodynamics", *Cognitive Science* Vol. 1, No. 2, pp. 192–215.

Newell, A., "AI and the Concept of Mind" in Schank, R. C. and K. Colby, *Computer Models of Thought and Language* (San Francisco 1975).

Simon, D. P. and Simon, H. A., "Individual Differences in Problem Solving in Physics," in: Siegler, R. S. (ed.) *What Develops*? (Hillsdale, N.J. 1979).

* * *

5. Deskriptive Psychologie und Phänomenologie

5. Descriptive Psychology and Phenomenology

THE DESCRIPTION OF MENTAL STATES

Aaron Ben-Zeev
University of Haifa

The distinction between different levels of description is of the utmost importance in order to understand the nature of mental states. Such understanding is found, by and large, in Aristotle's view. I first mention Aristotle's view of the four explanatory factors. Accordingly, I indicate the way we should describe mental states.

Aristotle assumes that everything can be explained in the light of four explanatory factors ("causes" as they are inadequately termed): material, formal, efficient and final. The *material cause* refers to a lower level of description than that of the explanandum, namely, the elements out of which something is made. The *formal cause* refers to a higher level—that into which something is made. This cause refers to the essential functions (or the general regularity) typical of the species in question. Discussing the species is on a higher level than discussing the individual animal. The *efficient cause* is that by which something is done, and the *final cause* is that for the sake of which something is done. Both of the last causes usually refer to the same level of description. The movement of one billiard ball is on the same level of description as the movement of another ball which is its efficient cause. Similarly, the mature tree is on the same level of description as the young tree (the former is the final cause of the latter).

The material and formal causes are not separate in time or space from the explanandum; they are different aspects of it. In the case of the efficient and final causes, it is not the level of description that differentiates them from the explanandum. Rather, they refer to events that are separate—in time or space—from the explanandum. While material and formal causes refer to what constitutes a certain thing, efficient and final causes refer to causes of, and reasons for, a certain process of change. The relations in the former are not causal relations in the usual sense, since there is no actual separation between the cause and the effect. They are, rather, relations of correlation, namely, a relation of support (the lower level entities support the higher level properties) or of realization (the higher level properties are realized in the lower level entities). The efficient cause is identified with what modern science usually regards as a cause. The scientific notion of causal relations refers to relations within a certain level of description. The relation between the explanandum and the final cause is close, I believe, to the modern notion of intentionality. The final cause is that which the activity is directed at; intentionality is the feature of being directed at.

In the light of the above explanatory framework, we may explain each phenomenon by referring to at least three levels of description: (A) A lower level of description—something like Aristotle's material cause; the relation between the entities of that level and the explanandum is that of *support.* (B) The same level of description—something like Aristotle's efficient and final causes; here the relations are those of *causality* or *intentionality*. (C) A higher level of description—something like Aristotle's formal cause; the relation here is that of *realization.* While causality is taken to be a relation between two different things, support and realization are relations of correlation between two different aspects of the same thing.

It should be noted that the "height" of a certain level refers to its degree of complexity. A state consisting of a certain structure of atoms is, in this sense, more complex than a single atom; hence, it is described on a higher level of description. The higher-level properties are dependent upon the lower-level ones, but not vice versa. Accordingly, the lower-level properties are more fundamental and less conditioned.

Let us now apply the above framework to two phenomena: a water molecule, and a mental state of fear. A water molecule may be described by referring to: (a) protons, neutrons, and

other elementary particles which constitute the molecule of water—this is a relation of support, that is a relation between a supportive basis and a supported state; (b) other water molecules—here we find relations of causality, namely, between an efficient cause and its effect; (c) properties of water as liquid, for example, its being wet, it pours, etc.—here we find the relation of realization, for instance the relation between a whole and its parts. The relation between the elementary particles constituting a water molecule and that molecule is different from the relation between two molecules of water. The latter relation is between two entities separated in time and space; the former relation is between two different aspects of the same thing. In a similar manner, the properties of water as liquid are not separate (in time or space) from the individual water molecules; they are just properties of a higher level of description. The relation between a mental state and a physiological structure is similar (I argue below) to the relation between a water molecule and its elementary particles, or to that between the liquidity of the water and the water molecules. It is a relation of realization. As Searle nicely puts it, "The liquidity of a bucket of water is not some extra juice secreted by H_2O molecules. When we describe the stuff as liquid we are just describing those very molecules at a higher level of description than that of the individual molecule.The liquidity, though not epiphenomenal, is realized in the molecular structure of the substance in question".[1] In this respect, higher-level properties may be a certain explanatory factor for lower level behavior. Thus, we may say that in the light of the water liquidity, further tilting of the pitcher is a reason for the individual water molecules to be in a different location. Similarly, we may describe physiological events by referring, among other things, to their mental correlates. The factor of one's will may be incorporated in the general explanation in this way. However, here also the relation between the higher-level properties and the lower-level ones is not that of causality. Rather, it is a correlation which is used as an explanatory short-cut.

One's state of fear may be described by referring to (a) a stimulation of a physiological mechanism, (b) a particular mental event, e.g., seeing one's boss or thinking about one's future promotion, and (c) the general character (structure) of the agent. The relation between physiological and mental events is not that between a cause and its effect, but rather a relation of support, or realization, between events from different levels of description. In our example, seeing one's boss (which is an efficient cause) and thinking about one's future promotion (which is a sort of final cause) are explanatory factors at the same level of description as that of the agent's state of fear, since all of them are particular mental states. It can be seen that though the final cause refers to a future event (a future promotion) it is not an occurrent, future event that finally influences the agent's state of fear; it is rather a present mental state, such as the agent's thought of the promotion that actually causes the fear. The reference to a future event has, nevertheless, some explanatory value, but it is not that of a retroactive, efficient cause. The formal cause in our example is the agent's general character (structure). Certainly, one with a different character would not be frightened at seeing one's boss or a future promotion. Here we do not refer to a particular mental state, but to a complex system of (dispositional and occurrent) mental states. The relation here is a relation of realization which exists between a whole and its parts. A whole is on a different level of description than that of its parts. We do not have a physiological event and a mental event; rather we can treat the same event as a physiological or a mental phenomenon.[2]

In the same vein, Aristotle argues that anger, like other affections, has a double character. A physicist would define anger as a boiling of the blood and hot stuff around the heart, while a dialectician would define anger as the desire for returning pain for pain (*De Anima*, 403a 29ff.). The desire for retaliation cannot be found in the boiling blood, though the boiling blood is a necessary supportive basis for that desire. In explaining the desire one should refer, as an efficient cause, to the pain suffered and not to the boiling blood.

By saying that properties belonging to two different levels of description express different aspects of the same things, I mean that what was separated, through abstraction, by our cognitive system when it used two different levels of description, may be found as the same thing in a level independent of those two levels (or in a level completely independent of our cognitive

system, i.e., something like the thing-in-itself). Identification, as abstraction and negation, is an act of the cognitive system and not a feature of the thing-in-itself (if one wants to use this notion). As Wittgenstein remarks "It is self-evident that identity is not a relation between objects".[3] In identification, something that was distinguished in abstraction is being reunited by the act of identification. In a sense, there is here cancelation of a previous act of the cognitive system (i.e., abstraction), and an attempt to return to a level less independent of the cognitive system. Clearly, the notion of identity used here is different from the pure logical notion of identity. Concerning the latter, Wittgenstein rightly argues that "Roughly speaking, to say of *two* things that they are identical is nonsense, and to say of *one* thing that it is identical with itself is to say nothing at all".[4] Actually, I do not speak here about pure identity, but rather about an act of identification. This act (of the cognitive system) is correlative to the act of abstraction (or to that of making distinctions). Maybe we should speak here not about identity but about a sort of non-causal correlation.

The view suggested here is somewhat similar to Spinoza's view of different sets of attributes of the same thing. The various levels of description are different hands of the same clock; therefore, one cannot speak about causal relations between them. Similarly, one cannot speak about causal relations between a whole and its parts. In both cases there is not the separation needed for causal relations, and hence we may speak merely about correlations. My view differs from Spinoza's by not assuming a complete symmetry between the mental and the physical levels of description. Mental states emerge only out of certain complex physical systems. Hence, whenever there is a mental state there is a physical state as well, but not vice versa. The structure of a physical (or biological) system is very important for determining the possible emergence of mental states. In this sense, matter matters. This is also Aristotle's stand. He criticizes his predecessors for connecting the soul to a body "without adding any specification of the reason for their union or of the bodily conditions required for it" (*De An.*, 407 b13). In Aristotle view everything has a form, but not every form is mental. My view is also close to Hartmann's and Ryle's views. Both speak about various levels (or categories) of descriptions. Both assume that the regularity of the lower level does not unequivocally determine the events of the higher level; it determines what kinds of events cannot take place, but it leaves a few possible options regarding the particular event that actually will take place.[5]

In the approach suggested here the relation between physiological and mental events is no longer so mysterious; similar relations can be found in other phenomena as well. One important difficulty of the old enigma remains in this approach too, i.e., why is a particular physiological mechanism responsible for a particular mental state? At this stage of our knowledge we are far from answering this question. In the water example, such a question can have a meaningful answer. It can be indicated why the properties of liquidity arise out of H_2O molecules only when they move at a certain speed (at different speeds they form gas or ice). This indicates that though the types of levels of descriptions are similar in the above phenomena, the gap between the levels in the case of mental states is still the largest. Nevertheless, it is possible that future empirical research may narrow this gap. For example, such research may show that the physiological mechanism underlying pain feeling operates in dangerous circumstances for the organism; in this sense pain is a clear warning signal for these circumstances. The kind of correlations here may be similar to those existing between properties of a field and properties of its components.

Using different levels of description seems to be an indispensable tool in everyday and scientifc explanation. I see no reason why should we not use this tool in describing and explaining mental states as well. Though this tool may seem to assume too obvious and simplistic an attitude, I strongly believe it can resolve many conceptual puzzles concerning the description of mental states. As Wittgenstein remarks, an important task of philosophy is that of reminders of the obvious.

ENDNOTES

[1] Searle, J., *Intentionality* (Cambridge 1983), p. 266.
[2] I apply this view to the explanation of perceptual states in my articles, "Toward a Different Approach to Perception", *International Philosophical Quarterly*, Vol. 23 (1983); and "The Kantian Revolution in Perception", *Journal for the Theory of Social Behavior*, Vol. 14 (1984).
[3] Wittgenstein, L., *Tractatus Logico-Philosophicus* (London 1961), 5.5301.
[4] *Ibid*, 5.5303.
[5] See, e.g., N. Hartmann, *New Ways of Ontology* (Westport 1975); G. Ryle, *The Concept of Mind* (London 1949).

* * *

DAS LEIB-SEELE-PROBLEM IN DER TRANSZENDENTALEN PHÄNOMENOLOGIE

Hans Köchler
Universität Innsbruck

Das Leib-Seele-Problem, wie wir es aus der metaphysischen Tradition kennen, gewinnt in der transzendentalen Phänomenologie eine neue Dimension, die sich von der spezifischen Formulierung des Subjekt-Begriffes herleitet. Wenn Husserl versucht, das personale Ich als „psychophysische Einheit" zu beschreiben,[1] so ist dies nicht im Sinne einer realistischen Ontologie zu verstehen, die eine Gleichrangigkeit naturwissenschaftlicher und geisteswissenschaftlicher Betrachtung bedeuten würde. Denn nach seiner transzendental-idealistischen Auffassung ist es geradezu sinnlos, „für die geistigen Eigenheiten eines Menschen eine psychophysische ‚kausale Erklärung' zu fordern als das, was das Nötige und einzig wissenschaftlich Nötige sei, um das zu leisten, was Naturerklärung in der physischen Sphäre leistet."[2] Die Aufgabe phänomenologischer Deskription des Leib-Seele-Bezuges muß mithin auf einer anderen Ebene liegen. In Husserls Definition ist „Seele" die Einheit der auf den niederen sinnlichen aufgebauten . . ., geistigen Vermögen'",[3] sie gewinnt ihre Realität dadurch, daß sie „als Einheit des seelischen Lebens verknüpft ist mit dem Leib als Einheit des leiblichen Seinsstromes, der seinerseits Glied der Natur ist."[4] Dieses „Ineinader von leiblichem und beseelendem Geschehen"[5] bedeutet, daß Seelisches niemals isoliert von der „Typik" eigener Leiblichkeit erfahren werden kann.[6] Die Mannigfaltigkeit dieser Interdependenz zu beschreiben, ist wesentliche Aufgabe phänomenologischer Deskription. In den „Ideen" bemüht sich Husserl, aufzuzeigen, „daß die ‚Natur' und der Leib, in ihrer Verflechtung mit diesem wieder die Seele, sich in Wechselbezogenheit aufeinander, in eins miteinander konstituieren."[7] Den Leib verteht er in diesem Sinne als „Umschlagstelle von geistiger Kausalität in Naturkausalität",[8] in ihm wandeln sich lediglich kausale Beziehungen in konditionale Beziehungen zwischen „Außenwelt" und „leiblich-seelischem Subjekt".[9] Die seelische Realität wird also durch den Leib konstituiert;[10] er ist gewissermaßen der „Nullpunkt" der Weltorientierung des Subjekts.[11] So ist auch letzlich das gesamte Bewußtsein des Menschen „durch seine hyletische Unterlage mit seinem Leib in gewisser Weise verbunden".[12] Mein Leib als „Nullkörper im absoluten Hier"[13] ist auch die Voraussetzung jeglicher Natur-Konstitution, die sich „in intentionaler Bezogenheit auf meine als wahrnehmend fungierende Leiblichkeit" vollzieht.[14]

Mit dieser Beschreibung der konkreten Leib-Seele-Beziehung ist jedoch – im Kontext der transzendentalen Phänomenologie – erst ein vordergründig-phänomenaler Problembestand in den Blick gekommen. Denn Husserl unterscheidet vom *realen seelischen Subjekt* – dem „identischen psychischen Wesen, das real verknüpft mit dem jeweiligen Menschen- und Tierleib das substantiell-reale Doppelwesen Mensch oder Tier, Animal, ausmacht"[15] – das reine oder *transzendentale Ich*, das der Unterscheidung von „Welt" und „Erfahrung", Leiblichem und Seelischem, als das ontologisch Grundsätzlichere vorausliegt.[16] Während er der „materiellen Welt" (innerhalb der als Natur konstituierten „objektiven Welt") eine Eigenständigkeit einräumt, aufgrund deren sie „keines Sukkurses anderer Realitäten bedarf",[17] postuliert er gleichzeitig, daß das *Bewußtsein* als solches grundsätzlich „*ohne eine Natur*" denkbar sei.[18] Husserl glaubt, daß das Bewußtsein, sobald es „rein gesetzt ist als es selbst in phänomenologischer . . . Reduktion",[19] seine apperzeptive Auffassung als seelische Zuständlichkeit ablegen und seine „empirische Einordnung" in die Natur-Kausalität überwinden könne.

Die „verweltlichende Apperzeption" mit ihrer empirischen Formulierung des Leib-Seele-Wechselbezuges ist gewissermaßen nur die erste Stufe eines transzendental-phänomenologischen Rückganges „auf mich als das universale, absolute, das transzendentale Ego."[20] Die

natürliche Einstellung", die uns die konkrete psychophysische Einheit des Menschen beschreiben läßt, wird in Husserls phänomenologischer Konstitutionsanalyse aufgehoben in der transzendentalen Einstellung, durch die das „absolute naturkonstituierende Bewußtsein" hervortritt,[21] auf das alle Natur – durch die von Husserl postulierte Korrelation zwischen Konstituierendem und Konstituiertem – relativ sein soll. Das Natürlich-Reale, dem auch der leibliche und seelische Aspekt der psychophysischen Einheit zuzuordnen sind, wird so in seiner ontologischen Bedeutung wieder entwertet: die reine transzendentale Subjektivität gilt als das ontologisch Absolute.[22] Die konkrete seelische Welterfahrung wird so als eine in der Monade des transzendentalen Ego „sich vollziehende Selbstobjektivierung derselben"[23] verstanden. Liegt der Seinssinn der Körperlichkeit – und mithin auch der Leiblichkeit, sofern sie der Naturwirklichkeit zugeordnet ist – in der Kausalität und der damit gegebenen Identifizierbarkeit und Unterscheidbarkeit als physische Individualität, so hat das (transzendentale) Ich bereits Individualität „aus sich selbst",[24] sein Fürsichsein wird – nach Husserls Auffassung – in keiner Weise durch die psychophysische Einbindung in die Raumzeitlichkeit fundiert: für das transzendentale Ich „sind Raum und Zeit keine Prinzipien der Individuation".

Dies bedeutet nach Husserl, daß der „Geist" (das transzendentale Subjekt) immer nur bis zu einem gewissen Grade „naturalisiert" werden kann,[25] daß mithin seine eigenen Analysen der Konstitution des Ego als „psychophysischer Einheit"[26] nur relative Bedeutung haben, insofern sie das personale Ich als „unmittelbar waltend in meinem, dem einzigen Leib, unmittelbar auch hineinwirkend in die primordiale Umwelt",[27] aufweisen, ohne aber in dieser „Manifestationsform" seine ontologische Bedeutung zu erblicken. Denn Subjekte können nach Husserl „nicht darin aufgehen, Natur zu sein, da dann das fehlen würde, was der Natur Sinn gibt".[28] Natur ist für ihn ein Feld „durchgängiger Relativitäten", die zum Geist als dem alle Relativitäten „tragenden" Absoluten hingeordnet sind.[29] Die Absolutheit des Geistes besteht also *trotz* der Bestimmtheit durch seine Umwelt,[30] wie sie Husserl ja auch in detaillierten phänomenologischen Analysen faktisch herauszuarbeiten suchte. Das Subjekt „waltet" in seinem Leib,[31] es besitzt ontologische Priorität. So hat für Husserl die eigentliche, konkrete psychophysische Einheit ihren objektiven Sinn ausschließlich „als Einheit des Waltens einer Subjektivität" im Leibkörper, der dadurch „subjektiv fungierendes Organ" wird.[32] „Geistigkeit (transzendentale Subjektivität) besteht nach dieser in eigentlichem Sinne idealistischen Auffassung bereits *vor* der raumzeitlichen Welt; davon ist die Dimension des „Seelischen" unterschieden, das die Weise bezeichnet, „wie Geistigkeit in der Raumwelt Lokalität, gewissermaßen ihre Verräumlichung gewinnt, und in eins mit ihrer körperlichen Unterlage *Realität*."[33]

II

An der Beschreibung des Vorganges des Todes wird nun konkret die idealistische Grundhaltung deutlich, die Husserls phänomenologische Beschreibung des Leib-Seele-Problems durchformt: in der „natürlichen Weltbetrachtung" ist die konkrete Leiblichkeit Bedingung der Möglichkeit „seelischen Seins"; dieses wird zu einem „Nichts", sobald die apperzeptiven Voraussetzungen seelischer Erfahrung – im Tod – aufgehoben werden.[34] Diese Interpretation des Todes als eines „realen Vorkommnisses in der Welt"[35] bezieht sich jedoch nur auf die vordergründige Ebene des naiven Weltvollzuges, der noch nicht zur Bewußtmachung der transzendentalen Subjektivität vorgedrungen ist. Auf der Ebene des reinen Subjekts gewinnt der Begriff der „Unsterblichkeit" eine neue Bedeutung, nicht als „Ablösung der Seele zu einem eigenen Realen" innerhalb der Welt,[36] sondern als Inbegriff der Nichtrückführbarkeit transzendentaler Subjektivität auf einen Kontext „weltlicher" Kausalitäten. Die „weltliche" Betrachtung des Todes als Vernichtung der Seele *in* der Welt[37] muß so nicht in einem Gegensatz zur transzendental-idealistischen Betrachtung stehen: „Die Unsterblichkeitslehre müßte also, wenn sie dem Sinn der Welt, wie er durch die universale objektive Erfahrung festgelegt ist, nicht widersprechen soll, eine ganz andere Bedeutung haben und kann sie in der Tat

haben, wenn es wahr ist, daß die natürliche Weltbetrachtung, die alles natürlichen und weltkundlichen Lebens, nicht das letzte Wort behalten muß und vielleicht behalten darf."[38] Husserl will in seiner Phänomenologie als transzendentalem Idealismus gerade zeigen, daß die Gesamtheit möglicher Gegenstände objektiver Erfahrung (was wir als „Welt" im natürlichen Sinne bezeichnen) nicht als im absoluten Sinne Seiendes gelten darf, daß das Absolute, „das die Welt schon voraussetzt", vielmehr der „Geist" (i.e. das transzendentale Subjekt) ist, jedoch nicht in seiner leiblich-gegenständlichen Konkretionsform – als „verweltlichter Geist" –, d.h. nicht als „Seele".[39]

Damit wird nochmals deutlich, in welch strikter Weise die transzendentale Phänomenologie Husserls zwischen Seelischem und Geistigem unterscheidet: der Seele als *konkreter* Subjektivität, die sich in realer Einheit mit dem „Leibkörper" manifestiert,[40] liegt die *reine* (transzendentale) Subjektivität als deren ontologische Bedingung voraus, welche die „weltlichen" Konkretionsformen von „Leib" und „Seele" als Konstitutionsformen umgreift. Damit geht Husserl über seinen ursprünglich realistischen Ansatz, wie er sich in der Analyse des Bezuges von Leib und Seele auch in den „Logischen Untersuchungen" manifestierte, hinaus. Während er in den „Logischen Untersuchungen" der Relativität der Existenz der Welt noch eine Relativität der Existenz des Ich zuordnete,[41] dergestalt eine reale psychophysische Wechselwirkung postulierend, und den Unterschied zwischen den Bereichen des Physischen und des Psychischen – in genuin phänomenologischer Weise – vorerst einmal in der verschiedenen Art ihrer *Gegebenheit* erblickte, wird ihm dieser in den „Ideen" zu einem ontologischen Gegensatz, der in der konkreten psychophysischen Einheit natürlicher Welterfahrung nicht überwunden, nur verdeckt werden kann. Während in den „Logischen Untersuchungen" diese psychophysische Einheit noch als Grundlage des Bewußtseins angenommen wurde, ist nach der idealistischen Wende in den „Ideen" das Bewußtsein etwas in sich Selbständiges, Bewußtseinserfahrungen können nur noch eine Einheit *miteinander* bilden.[42] Dies führt dazu, daß auch die konkrete Leiblichkeit des Menschen relatitiviert wird, da alles Wahrgenommene letzlich seinen Seinssinn als „intentionales Korrelat des Bewußtseins" erhält und mithin auch der eigene Leib als von mir Wahrgenommenes (Objekt meiner Wahrnehmung) in konstitutiver Relation zu meinem Bewußtsein steht. Die konkrete Einheit (und Wechselwirkung) von Leib und Seele erweist sich somit als Schein der „natürlichen" Einstellung. Auch in seiner Theorie der Leibeskonstitution[43] hat Husserl diese im letzten idealistische Ontologie nicht aufgeben; diese Theorie ist vielmehr nur eine Ergänzung seiner Konstitutionsanalysen auf der Grundlage derselben ontologischen Voraussetzungen, nicht eine Revision dieser Voraussetzungen selbst.

Auch für die Analyse der Konstitutionsform der Leiblichkeit gilt, was Husserl generell für die phänomenologische Analyse von „Gegenständlichkeiten" fordert: „Durch Rückgang auf das absolute Bewußtsein und die in ihm zu verfolgenden gesamten Wesenszusammenhänge sind allererst die sinngemäßen Relativitäten der betreffenden Gegenständlichkeiten der einen und anderen Einstellungen und ihre wechselseitigen Wesensbeziehungen zu verstehen."[44]

Hier nun eröffnet sich ein neuer Sinn von „Transzendenz", von dem Husserl nicht ausdrücklich spricht, der aber bezeichnend für jede idealistisch-transzendentale Innenwendung ist: als letzte Möglichkeitsbedingung wird ein *absolutes* transzendentales Ich postuliert, das jeglicher empirischen Erfahrung des Bewußtseins als deren Bedingung immer schon vorausliegt und so – als zentrierender „Ichpol"[45] – das oberste synthetische Einheitsprinzip aller Erfahrung ausmacht. Dieses Ich ist selbst – so Husserl – „identische Einheit",[46] die sich in allen einzelnen cogitationes manifestiert. Es ist prinzipiell meta-empirisch, auch in iterierter Reflexion nicht einholbar. Es „ist das identische Subjekt der Funktion in allen Akten desselben Bewußtseinsstroms, es ist das Ausstrahlungszentrum, bzw. Einstrahlungszentrum alles Bewußtseinsleben, aller Affektionen und Aktionen".[47] Als Bedingung jeglichen Wandels des empirischen Ich ist es also gerade das Unwandelbare, Ab-solute (d.h. vom Werden Losgelöste), jenseits gegenständlicher Relationen; denn – in Husserls Worten – „*jedes cogito mit allen seinen Bestandstücken entsteht oder vergeht im Fluß der Erlebnisse. Aber das reine Subjekt entsteht nicht und vergeht nicht*, obwohl es in seiner Art ‚auftritt' und wieder ‚abtritt'".[48]

Es ist – gegenüber den einzelnen cogito's – je schon transzendent, d.h. kein „reelles Moment derselben".[49] Es ist inkommensurabel mit dem empirischen Ich, es geht nur in bestimmte Manifestationsformen ein, durch die sich dann ein empirisches Ich konstituiert.[50] In solcher Konstitution einzelner cogitationes vollzieht sich seine Selbsterfassung.

III

Das „reine" (transzendentale) Subjekt ist somit aus dem „empirischen" Kontext des weltlichen Leib-Seele-Bezuges herausgehoben, es ist dessen ontologisches Fundament. Der Aufweis naturhafter Abhängigkeiten dieses als „Geistiges" verstandenen Transzendentalen – etwa mit dem Ziel naturwissenschaftlich eindeutiger Bestimmung des Geistigen – ist für Husserl daher undenkbar.[51] Damit wird die von ihm selbst durchgeführte detaillierte Schilderung der psychophysischen Wechselbeziehungen in ihrem philosophischen Stellenwert relativiert: seine Analyse des psychophysischen Leibes als „primordialen" Faktors bei der Konstitution der objektiven (raum-zeitlichen) Welt[52] wird durch das Diktum der „Irrelativität" der Subjektivität in ihrem realistischen Charakter wieder negiert. Der Leib als derjenige Körper, in welchem die Subjektivität mit der Realität eines diese transzendierenden Ansich vermittelt wird, wird so schon von vornherein als ontologisch eigenständige Kategorie aufgehoben. Der Unterschied zwischen Seelischem und Materiellem – den Husserl als „Realitätsüberschuß über das bloße physische Ding" betrachtet,[53] wobei im Leib beide Realitätsdimensionen psychophysisch vereint sind und Seele dergestalt als eine „bloße Schicht realer Vorkommnisse an Leibern"[54] verstanden wird –, dieser Unterschied erscheint nunmehr als vordergründig, als nicht mehr konstitutiv für eine ontologische Wertung der Leib-Seele-Problematik. Der Leib, den er – in seiner Verflochtenheit mit dem seelischen Subjekt – als das „subjektive Objekt" bezeichnet,[55] und die Seele, die zusammen die „beseelte Realität" des personalen Ich ausmachen, welche nach beiden Seiten als Substrat kausaler Eigenschaften fungiert,[56] sind lediglich verschiedene Manifestationsformen der transzendentalen Subjektivität, in denen sie sich als „Mundanität" konstituiert. Der ontologische Idealismus, wie wir ihn an Husserls transzendentaler Phänomenologie zu explizieren suchten, läßt den Leib-Seele-Gegensatz nicht mehr als reelle Unterscheidung von seinsmäßig Verschiedenem gelten. (Damit hängt auch eine gewisse Künstlichkeit in der Analyse der Erfahrung des Fremdpsychischen zusammen,[57] da der ontologische Idealismus Husserls letzlich eine solipsistische Fiktion bedeutet, die eine Begegnung mit der Personalität des Anderen nur über eine analogisierende Einfühlung innerhalb der eigenen, in sich geschlossenen Subjektivität möglich macht.)

Der ursprünglich realistische Ansatz der Analyse des Leib-Seele-Bezuges, wie er in den „Logischen Untersuchungen" noch deutlich wird, ist durch Husserls Glauben an das „vielleicht in sich absolut geschlossene eigene Wesen" des Seelischen[58] nicht nur in Frage gestellt, sondern aufgehoben worden. Auch wenn sich in der vorhin zitierten Formulierung eine gewisse Unsicherheit manifestiert, mit der auch die Ambivalenz und Unklarheit der Termini „Seele" und „Geist" zusammenhängt, die Husserl häufig mit einander überlagernden Bedeutungskomponenten verwendet, bleibt die Tatsache bestehen, daß die „induktiv wirksame Typik" des Seelischen,[59] d.h. seine Kausalität in Abhebung vom Physisch-Materiellen, und die „induktive Kausalität" des Physischen[60] nicht im ontologischen Sinne eigenständigen Realitätsbereichen entsprechen. Das einzige, was Realität konstituiert, ist das reine Bewußtsein (reine Subjekt). In der transzendentalen epoché wird auch der Leib – als ein von dieser Subjektivität konstituierter spezieller Gegenstand, der in einer Wechselwirkung mit dem empirischen Bewußtsein steht und dadurch vor anderen Gegenständen ausgezeichnet ist – reduziert auf seinen Charakter, Gegenstand reiner fungierender Intentionalität zu sein. Seine „Auszeichnung" ist nur noch eine unter konstituierten Gegenständen, nicht eine etwa von seiner „inneren" (ontologischen) Struktur abgeleitete. Die realistische Ursache-Wirkung-Beziehung wird damit geradewegs umgekehrt: nicht der (die Verbindung zum Materiellen herstellende) Leib bestimmt das Bewußtsein, sondern das Bewußtsein konstituiert den Leib als eine spe-

zielle Gegenständlichkeit innerhalb der Mundanität. Zwar ist der Leib als „Naturding" in die Kette der „realen" Kausalzusammenhänge eingegliedert und das empirische Bewußtsein über den Leib wiederum von diesen Kausalwirkungen bestimmt – was Husserl auch in seinen phänomenologischen Deskriptionen in den „Ideen" darlegt –, aber die „Realität" dieser Zusammenhänge muß eben im Lichte seiner transzendentalen Phänomenologie reinterpretiert werden als „Natur" im *transzendentalen* Sinne, d.h. als eine im reinen Bewußtsein „durch immanente Zusammenhänge motivierte intentionale Einheit".[61] In seiner transzendentalen Reduktion, die nicht nur ein methodischer Schritt zur besseren Beschreibung der Phänomene, sondern eine ontologische Setzung ist,[62] erblickt Husserl den eigentlichen Schritt zur universalen und endgültigen Formulierung der Realitätsproblematik, wenn er gegenüber einer realistisch argumentierenden Kritik betont: „Wir haben eigentlich nichts verloren, aber das gesamte absolute Sein gewonnen, das, recht verstanden, alle weltlichen Transzendenzen als intentionales Korrelat der ideell zu verwirklichenden einstimmig fortzuführenden Akte habitueller Geltung in sich birgt, sie in sich ‚konstituiert'."[63] Eine in diesem Sinne idealistische Position läßt die Nivellierung der Leib-Seele-Problematik auf die Ebene von mundanen „Konstitutionsformen" einer reinen Subjektivität offenbar nicht als Verlust erscheinen. Gleichwohl vermag auch die auf der transzendentalen Reduktion (wie Husserl sie versteht) basierende Phänomenologie die stets von neuem drängende Frage nach der real-ontologischen Eigenständigkeit des Bewußtseins gegenüber dem materiellen Sein – mit all den existenziellen Implikationen, die sich aus ihrer Beantwortung ergeben mögen – nicht aus der Welt zu schaffen.

ANMERKUNGEN

1 *Cartesianische Meditationen* (ed. Ströker) (Hamburg 1977), S. 100.
1 *Husserliana* (abgekürzt: Ha.), Bd. VI, S.479.
3 *Ha.* IV, S. 120.
4 a.a.O., S. 139.
5 *Ha.* IX, S.107.
6 a.a.O., S. 108.
7 *Ha.* V, S. 124.
8 *Ha.* IV, S. 286.
9 ebd.
10 Vgl. *Ha.* IV, S. 143ff.
11 a.a.O., S. 158.
12 a.a.O., S. 153.
13 *Cartesianische Meditationen*, a.a.O., S. 126.
14 a.a.O., S. 119.
15 *Ha.* IV, S. 120.
16 Allerdings scheint Husserl – in einem anderen Kontext – sich selbst zu widersprechen und diese strenge Unterscheidung wieder in Frage zu stellen, wenn er (in einer Kritik an Kant) betont, daß, „sowie wir diese transzendentale Subjektivität aber von der Seele unterscheiden," wir in ein „unverständlich Mythisches" geraten (*Ha.* VI, S. 120).
17 *Ha.* V, S. 117.
18 *Ha.* IV, S. 118.
19 ebd.
20 *Cartesianische Meditationen*, a.a.O., S.102.
21 Ha. IV, S.179.
22 Vgl. dazu auch die Arbeit des Verf.: *Die Subjekt-Objekt-Dialektik in der transzendentalen Phänomenologie* (Meisenheim a. G. 1974).
23 *Cartesianische Meditationen*, a.a.O., S. 134.
24 *Ha.* VI, S. 222.
25 *Ha.* IV, S. 297.
26 *Cartesianische Meditationen*, a.a.O., S. 113.
27 ebd.
28 *Ha.* IV, S. 297.

[29] ebd.
[30] ebd.
[31] *Ha.* IX, S. 395.
[32] ebd.
[33] a.a.O., S. 132.
[34] a.a.O., S. 109.
[35] ebd.
[36] ebd.
[37] ebd.
[38] ebd. – Vgl. auch Ms. K III 6 (Husserl-Archiv), S. 251a.
[39] ebd. – Zum Problem der terminologischen Abgrenzung der Begriffe „Leib", „Seele", „Geist" und der damit zusammenhängenden je verschiedenen Bedeutung des Leib-Seele-Dualismus vgl. Josef Seifert, *Das Leib-Seele-Problem in der gegenwärtigen philosophischen Diskussion*. Eine kritische Analyse (Darmstadt 1979), S. 126ff.
[40] *Ha.* IX, S. 393.
[41] Vgl. *Logische Untersuchungen I*, Bd. 1, Halle a. S., 2. Aufl. 1913, S. 121; Bd. 2, Halle a. S., 1901, S. 337.
[42] Zu dieser Deutung vgl. auch Theodor De Boer, *The Development of Husserl's Thought* (The Hague/Boston/London 1978), S. 227ff.
[43] Vgl. *Ha.* IV, S. 157ff; *Ha.* III, S. 117ff.
[44] *Ha.* IV, S. 180.
[45] Vgl. *Ha.* IV, S. 105.
[46] *Ha.* V, S. 113.
[47] *Ha.* IV, S. 105.
[48] a.a.O., S. 103.
[49] a.a.O., S. 102 (im Original Singular).
[50] *Ha.* V, S. 114.
[51] *Ha.* IV, S. 297.
[52] *Cartesianische Meditationen*, a.a.O., S. 137.
[53] *Ha.* IV, S. 176.
[54] a.a.O., S. 175.
[55] *Ha.* V, S. 124.
[56] *Ha.* IX, S. 133. – Zu dieser das personale Ich bestimmenden Interdependenz und psychophysischen Einheit vgl. auch: *Cartesianische Meditationen*, a.a.O., S. 100.
[57] Vgl. *Cartesianische Meditationen*, a.a.O., S. 109ff.
[58] *Ha.* IX, S. 135.
[59] a.a.O., S. 138.
[60] a.a.O., S. 136.
[61] a.a.O., S. 120.
[62] Vgl. die Arbeit des Verf.: *Die Subjekt-Objekt-Dialektik in der transzendentalen Phänomenologie* (Meisenheim a. G. 1974), S. 79ff.
[63] *Ha.* III, S. 119.

* * *

THE UNIQUENESS-CLAIM ON PLEASURE AND ITS EMPIRICAL SIGNIFICANCE

Rainer Stuhlmann-Laeisz
Universität Göttingen

Is pleasure One or Many? This initial question from *Plato's Philebos* bears considerable import to ethical hedonism. Let us call the possible answers to the question the One-claim and the Many-claim on pleasure, respectively. In the first part of this paper I shall offer various meanings of these claims and have a look at their import to hedonism, in the second part I shall give empirical interpretations of the different formulations. However, I shall not decide the battle, since once the One/Many-question has been turned into an empirical one, this will not be a philosopher's job.

A. There are various activities or experiences one can get pleasure from. Someone might enjoy leading an excessive life, whereas someone else gets pleasure from sobriety. Especially examples like these made Plato ask the question.[1] Let us broach the problem in the following way. I presuppose that with every experience e giving pleasure to a person X there is associated a personal state F, e.g. a feeling, being the pleasure which X gets from e. On this basis we get a variety of One/Many-questions concerning pleasure, e.g. regarding only one person X we may ask whether (with all the experiences pleasant for X) F is the same rather than variable, regarding different persons—possibly all human beings—one can ask whether or not F is the same with all persons—and so on.—Putting things more exactly and distinguishing the case of only person from the case involving different persons we in detail get the following claims (the variables 'F' and 'e' running over personal states and experiences or activities respectively).

(i) The one-person-case. Here we have as the One-claim on pleasure:

> (One-pers-One-claim) There is an F such that for all e: if the person X enjoys e, then F is the pleasure which X gets from e.

The corresponding Many-claim would be the following:

> (One-pers-Many-claim) There is e_1 and e_2 such that the person X enjoys e_1 and enjoys e_2, but the pleasure F_1 which X gets from e_1 is different from the pleasure F_2 which X gets from e_2.

This is, in the one-person-case, the weakest form of the Many-claim, since it calls only for two experiences or activities yielding different pleasures. The One-claim mentioned above, on the other hand, is a very strong one, since it implies that pleasure is unique for the person X: there is exactly one personal state F which is *the pleasure got by X from any pleasant experience*. This uniqueness is due to the fact that the phrase 'F is the pleasure which X gets from e' expresses a functional identity: F = Pl (X,e), 'Pl' denoting a function which attaches to each person X and every experience e pleasant to X the pleasure F got by X from e.—Note that with this interpretation the One-claim is contradictory to the Many-claim, thus they can neither be both true nor both false. The One-claim, however, is only contrary to those tightenings of the Many-claim which call for more than two different pleasures, the strongest of these tightenings being the claim:

(Strong One-pers-Many-claim) For any two experiences e_1 and e_2: if the person X enjoys both, e_1 and e_2, then the pleasure which X gets from e_1 is different from the pleasure which X gets from e_2.

This is to say that for X there are as many pleasures as there are pleasant experiences. If this claim is true, one might identify any pleasant experience or activity with the pleasure caused by it.

(ii) The different-person-case (I now use 'X' as a variable ranging over persons). Here we can distinguish two One-claims and at least three Many-claims:

(Diff-pers-One-claim a) For every e there is an F such that for all X: if X enjoys e, then F is the pleasure which X gets from e.

(Diff-pers-One-claim b) There is an F such that for every e and X: if X enjoys e, then F is the pleasure which X gets from e.

Claim (a) says that with every pleasant activity or experience e there is associated a unique personal state F being the pleasure all persons get from e. E.g. for all people who like drinking wine tipsiness might be the pleasure got by this activity. However, (a) allows what (b) denies, i.e. that pleasures vary with pleasant activities, thus (b) claims that pleasure is unique for all persons and all pleasant experiences.

Let us handle the different-persons-Many-claim in a less formal way. We first get two Many-claims as the negations of the One-claims. Claim (a') would say that at least one pleasant activity causes different pleasures to different persons, e.g., not all people might feel tipsiness as the pleasure got from drinking wine, but some might enjoy the pleasure of good taste. (b') would be the Many-claim in its weakest form, which speaking states that there are at least two pleasures, whereas the strongest claim would maintain that there are as many pleasures as there are connections of persons with pleasant experiences.

So far we h ave got different formulations of the two claims rendering precise their different meanings. Let us now have a look at the import these claims have to ethical hedonism. As is well-known, hedonistic theories hold the only intrinsic goods to be pleasant personal states, these and only these being desirable in themselves. In consequence, extrinsic goods are valued according to their contribution to those states, and these extrinsic goods are morally recommended to be chosen (as acts or states of the world to be realized or something like that). Well then, if the Diff-pers-One-claim b (the strong version) is true, hedonism—whether itself true or not—is apt to provide universal moral rules. If not, and if the strongest Many-claim is true, hedonism is likely not to be apt to provide such rules. The weaker formulation (a') would yield difficulties too, not so, possibly, the weakest claim (b').—

B. Let us now try to make things testable by giving empirical interpretations of the various One- and Many-claims. I shall do this by giving empirical content to the statements 'X enjoys e' and 'F is the pleasure got by X from e' and by rendering more precise the range of the variable 'F', so far only described as the class of personal states.—I propose the following decision-criterion for the truth or falsehood of statements 'X enjoys e':

X enjoys an experience e iff he, being in a position to choose between two states of affairs which—other things being equal—differ to the effect that in one of them he has the experience e whereas in the other one he has not, he,in the long run, would more often choose to have e rather than not to have e.

This criterion makes the question whether X enjoys an experience e or not decidable by observation of X's behaviour. If, e.g., e is the activity to play tennis, we can refute the claim that X enjoys e by pointing out a sufficiently large number of situations where X did not play tennis,

though he–other things left equal–could have done.–Let us call this criterion the *Sidgwick-criterion*,[2] and let us say that a person X satisfies the Sidgwick-criterion with respect to an experience e–abbreviated 'Sidg (X,e)'–if X behaves as described above. To make the statements 'F is the pleasure which X gets from e' empirically significant, I use the fact that in a certain sense one cannot be mistaken about how one feels, i.e. if a person X truly says that he has a certain feeling then X has this feeling. This is not to say that one cannot be mistaken in one's personal situation as a whole, this is, of course, the case as e.g. psychoanalysis shows. It is only to say that one cannot deceive oneself about one's momentary feeling, as well as one cannot be wrong about one's momentary sensation: if I hear a telephone ring, then in fact, I hear this ring–be there a telephone call or not. Of course, one can deceive other persons in this respect. Therefore, let us presuppose that a sufficiently large number of persons answers truly if asked whether they feel in a certain way. Under this assumption we can decide by observing X's answer to the corresponding question whether or not he feels that way.–Now I shall say that X is in the s-state if he affirms the question 'Do you feel s?', the variable 's' here running over expressions usable to form statements 'I feel s'. Of course, with this interpretation one can be in different states at the same time, e.g. if one feels cold and hungry; thus I now regard classes of states as candidates for pleasures to be got from experiences. Well then, for every X and e there is a unique class of states which accompany all occurrences of e to X (in some cases it might be empty, though). Denoting this class by 'Occ(X,e)' I replace the phrase 'the pleasure which X gets from e' by the notion 'Occ(X,e)', and I substitute the identity-sign '=' for the word 'is' within the statements 'F is the pleasure . . .'. The variable 'F' runs now over classes of states in the sense described.–Using these substitutes and the Sidgwick-criterion to replace the expression 'X enjoys e' we can turn our various claims into empirically significant statements. Take for example the strong Diff-pers-One-claim b. Here we get the formulation:

There is an F such that for every X ande: if Sidg(X,e), then F = Occ(X,e).

However, we must carefully watch the following possible outcome of this claim's test: It might turn out that the question 'Do you feel pleasant?' is the only one of this type which all people affirm if they are enjoying something. In this case the class F consisting of the pleasant-state alone would satisfy the strong claim above. But for sure, this would not be an honest proof. Therefore, I put the following constraint on F: I call F *neutrally describable*, if the conjuction of expressions constituting the states within F is definable independent of the notion of pleasure. If, e.g., F contains only the cold- and the hungry-state, then F is neutrally describable, but if F contains the happy-state, it might be not. Thus I finally add to each claim the requirement that the respective classes F or Occ(X,e) be non-empty and neutrally describable.–With this interpretation the weak Diff-pers-One-Claim reads:

> For every e there is a neutrally describable class of states, F, such that for all X: if Sidg(X,e), then F = Occ(X,e).

Perhaps, this claim is a good candidate for a true empirical statement.

ENDNOTES

[1] Vide *Philebos* 12 c 1 – d 7.
[2] I have derived the criterion from a definition of pleasure by Henry Sidgwick; vide: *The Methods of Ethics*. Book II, Chap. III, § 1.

* * *

INTENTION UND ERFÜLLUNG

Frank W. Liedtke
Universität Düsseldorf

Der Begriff der Erfüllung spielt in der Theorie der Intentionalität und in der Sprechakttheorie eine zentrale Rolle – wenn auch nicht seine einzige Erfüllung kann ausgesagt werden von intentionalen Akten, die sich auf Sachverhalte beziehen; oder von sprachlichen Äußerungen, mit denen Sprecher sich auf Zustände in der Welt beziehen; oder sie wird als Relation zwischen einer Aussagefunktion und einem Gegenstand aufgefaßt. Erfüllung ist das Resultat eines solchen In-Beziehung-Setzens, das in der Übereinstimmung der in Beziehung gesetzten Entitäten besteht. Daß dies ungefähr das allgemeine Schema ist dafür, wie der Erfüllungsbegriff verwendet wird, möchte ich im folgenden aufzeigen.

Umgangssprachlich wird Erfüllung von Erwartungen oder von Träumen ausgesagt, und man kann auch sagen, daß jemand seine Pflicht oder die ihm auferlegten Normen erfüllt. Ich werde hier die philosophische oder wissenschaftssprachliche Verwendung diskutieren, die sich von der umgangssprachlichen z.B. dadurch unterscheidet, daß sie viel weiter ist. Ich kann nur einige mir typisch erscheinende Verwendungsweisen aufzeigen, ohne die Heterogenität dieser Verwendungen verharmlosen zu wollen. Husserl verwendet den Begriff vielfältig. Ich möchte den Erfüllungsbegriff in seiner V. Logischen Untersuchung, die über intentionale Erlebnisse und ihre Inhalte geht, skizzieren. In dieser Untersuchung beschäftigt Husserl die Frage, worin die Akte des Bedeutens bestehen, und dies ist Teil seiner Frage nach dem Ursprung des Begriffes ‚Bedeutung'. Diese Frage wiederum geht aus seinem Interesse an einem Bedeutungsbegriff für sprachliche Ausdrücke hervor, der dann für die Aufstellung einer reinen Logik relevant wird. Die Akte des Bedeutens, oder: die bedeutungsverleihenden Akte, gründen auf dem meinenden Gerichtetsein auf einen Gegenstand, also auf intentionalen Erlebnissen. Dieses Begründungsverhältnis von intentionalen Erlebnissen und Bedeutungen sprachlicher Ausdrücke thematisiert Husserl in der Einleitung zur V. Logischen Untersuchung, und dies, indem er den Begriff der Erfüllung benutzt.

> Die Bedeutungen sollen in Aktintentionen liegen, die zur Anschauung in gewisse Beziehung treten können. Wir sprachen mehrfach von der Erfüllung der Bedeutungsintention durch korrespondierende Anschauung und daß die höchste Form dieser Erfüllung in der Evidenz gegeben sei. Es erwächst also die Aufgabe, dieses merkwürdige phänomenale Verhältnis zu beschreiben und seine logische Rolle zu bestimmen, d.h. die in ihm gründenden Erkenntnisbegriffe zu klären.[1]

Intentionale Erlebnisse oder Akte faßt Husserl als konstitutive Momente der individuellen Psyche auf, deren Besonderheit darin besteht, daß sie sich „in der Weise der Meinung oder in irgendeiner analogen Weise auf ein Gegenständliches beziehen."[2] Der intentionale Akt wird – möglicherweise – erfüllt durch den Gegenstand, auf den er abzielt. Husserl schreibt: „Im Bilde entspricht der Tätigkeit des Abzielens als Korrelat diejenige des Erzielens (das Abschießen und Treffen). Genauso entsprechen gewissen Akten als „Intentionen" (z.B. Urteils-, Begehrungsintentionen) andere Akte als „Erzielungen" oder „Erfüllungen".[3] Ausdrücklich analogisiert er das Verhältnis von Erfüllendem und Erfülltem dem Verhältnis von einem zustimmenden Urteil (Erfüllendes) auf eine Frage (Erfülltes) im Gespräch.[4] Von diesem diskursiven Modell aus reflektiert Husserl die Anwendung des Erfüllungsbegriffs auf intentionale Akte, also Vermutungen, Erwartungen, Hoffnungen, Wünsche. Daß das, was ich intendierend wünsche, eintritt, merke ich dadurch, daß Wunsch und Erwünschtes eine „Einheit im charakteristischen Erfüllungsbewußtsein" bilden.[5] Dieses Erfüllungsbewußtsein koordiniert den Wunsch, daß S P sei, und das Erfahren, S sei P. Das Erfahren, daß S P sei, ist urteilsmäs-

sig. Erfüllung von intentionalen Akten erfahren wir also immer in einem ganz bestimmten Modus – dem des Urteils. So führt die Erfüllung von Hoffnung, Wunsch, Erwartung etc. zu einem Übergang zu einem anderen intentionalen Akt: dem des Urteils. Nur das Urteil erfüllt sich in einem qualitativ identischen Modus: eben dem des Urteils. Wir können so eine wichtige Asymmetrie zwischen intentionalen Akten wie Hoffnungen und dem intentionalen Akt des Urteils feststellen: in einem Fall qualitative Verschiedenheit, im anderen Falle qualitative Identität des erfüllten und des erfüllenden Aktes.

Von dieser fragmentarischen Skizzierung des Husserlschen Erfüllungsbegriffs möchte ich zu einer Verwendung des Erfüllungsbegriffs übergehen, die von einer ähnlichen Asymmetrie gekennzeichnet ist. Der Begriff der Erfüllung hat nicht nur in der Aussagenlogik, sondern seit den dreißiger Jahren auch in der Logik normativer Ausdrücke – der deontischen Logik – einen festen Platz. Stellen Aussagen, sofern sie erfüllt sind, das Bestehen eines Sachverhalts fest und befinden sich so in Korrespondenz mit der Wirklichkeit, so ist Funktion von Befehlen, das ‚Herstellen' eines Sachverhalts zu fordern, der sich dann in Korrespondenz mit dem Inhalt des Befehls befindet und ihn so erfüllt. Der Rechtsphilosoph van Loon nennt dies treffend „correspondence – theory of oughtness".[6] Eine frühe Version dieser Strategie ist in einem Aufsatz von Hofstädter / Mc Kinsey von 1939 zu finden, einem locus classicus für die Korrespondenztheorie des Sollens. Die Autoren sehen der Einfachheit halber Imperativsätze als Sätze an, die aus Behauptungssätzen durch Hinzufügen eines Operators „!" gebildet sind. Die Bedingung für die Erfülltheit des Imperativs ist die Wahrheit des Behauptungssatzes, aus dem der Imperativsatz ja als abgeleitet gedacht wird. „Roughly, we understand an imperative to be satisfied if what is commanded is the case. Thus the fiat „Let the door be closed!" is satisfied if the door is closed."[7] Daß die Tür zu ist, muß behauptet werden können. Auch hier existiert also das Prinzip des Moduswechsels. Der Imperativ ist erfüllt, wenn festgestellt werden kann, daß das der Fall ist, was befohlen wurde. Das Prinzip der Rückführung von Imperativen auf Aussagen schildert N. Rescher treffend:

> The semantical theory of validity in command inference, although distinctive and in some ways sui generis, is not primitive and self-subsistent, but should be approached from the direction of assertoric logic via the bridging–link of statements of command–termination (or command ‚satisfaction' in some cognate sense).[8]

Eine weitere Theorie, die den hier aufgezeigten Traditionslinien mindestens insofern affin ist, als ihre grundlegenden Begriffe ebenfalls über das Prinzip der Erfüllung durch Korrespondenz definiert werden, ist John R. Searles Theorie der Intentionalität.[9] Er faßt Intentionalität als eine Eigenschaft vieler mentaler Zustände und Ereignisse auf, durch die sie auf Objekte oder Sachverhalte in der Welt gerichtet sind.[10] Die Beziehung der Theorie der Intentionalität zu seiner früheren Sprechakttheorie ist sehr eng.

Ein Sprechakt ist ja unter anderem definiert als der Ausdruck eines bestimmten intentionalen Zustands, und dieser ist umgekehrt der Inhalt der Aufrichtigkeitsbedingung des Sprechakts. In seinem Buch über Intentionalität expliziert Searle diese Analogie zwischen Sprechakten und intentionalen Zuständen. Zunächst leistet er dies über eine Angleichung der formalen Struktur. Sprechakte bestehen aus der illokutionären Kraft und einem propositionalen Gehalt, haben also die Form „F(p)"; intentionale Zustände werden aufgeteilt in einen repräsentativen Gehalt und einen Modus, in dem dieser Gehalt repräsentiert wird, haben also die Form „S(r)".[11] Aus dieser formalen Analogie ergibt sich die inhaltliche. Sie ergibt sich durch die für beide Begriffe – intentionaler Zustand wie Sprechakt – verbindliche Instanz der Erfüllungsbedingung, die jeweils in einem bestimmten Modus repräsentiert wird. Searle, der die Vagheit der Kategorie der Repräsentation zwar konzediert, aber ihretwegen nicht auf den Begriff verzichten will, schreibt:

> Exploiting this vagueness we can say that Intentional states with a propositional content and a direction of fit represent their various conditions of satisfaction in the same sense that speech acts with a propositional content and a direction of fit represent their condition of satisfaction.[12]

Der Begriff der Repräsentation ist explizit nicht im Sinne von Abbildung gemeint. Zu sagen, daß ein intentionaler Zustand wie ‚Glaube' eine Repräsentation sei, heißt nichts anderes, als ihm Eigenschaften zuzuschreiben, die auch Sprechakte besitzen, also vor allem: daß sein propositionaler Gehalt eine Reihe von Erfüllungsbedingungen determiniert. Sie sind die Begründungsinstanz für Sprechakte wie für intentionale Zustände, und das heißt für letztere: „. . . that the key to understanding representation is conditions of satisfaction."[13] Durch die Analogisierung mit Sprechakten, aber auch durch die Art der Theorie über Intentionalität wird deutlich, daß Erfüllung von intentionalen Zuständen – wie Erfüllung von Sprechakten – etwas mit der Wahrheit des repräsentierten propositionalen Gehalts zu tun hat. Und eine Entscheidung darüber kann wohl nicht anders zustande kommen als „urteilsmäßig", was uns zu der Vermutung führt, daß sich diese Theorie auch in der Hinsicht der Asymmetrie intentionaler Zustände in Bezug auf Erfüllung in der aufgezeigten Tradition befindet. Ich möchte nun Searles Ansatz und die Problematik seines Erfüllungsbegriffs durch Hinweis auf einige seiner Beispiele illustrieren, ohne auf seine Analogisierung von intentionalen Zuständen und Sprechakten näher einzugehen. Man kann hinsichtlich ihrer Erfüllung mindestens drei Klassen intentionaler Zustände unterscheiden. 1. Zustände, die ich nur haben kann, solange ich nicht weiß, ob sie erfüllt sind (Hoffen, Zweifeln, Erwarten). 2. Intentionale Zustände, die ich nur haben kann, solange ich glaube, daß sie erfüllt sind (Stolz, Scham, Enttäuschung). 3. Zustände, die keiner von beiden Restriktionen unterliegen. Man mag diese Unterteilung sehr streng finden, doch ich glaube, daß man sie machen kann, ohne das Haben irgendwelcher Intentionen verbieten zu müssen. In allen drei Fällen hat die Kenntnis der Erfüllungsbedingungen einen je unterschiedlichen Status, und man sollte dieser Differenz in einer Theorie der Intentionalität einen prominenteren Platz einräumen. Ein zweites Problem in Zusammenhang mit Searles Annahmen über Intention und Erfüllung läßt sich durch eine Bemerkung Wittgensteins demonstrieren:

> Der Wunsch scheint schon zu wissen, was ihn erfüllen wird, oder würde; der Satz, der Gedanke, was ihn wahr macht, auch wenn es gar nicht da ist! Woher dieses *Bestimmen* dessen, was noch nicht da ist? Dieses despotische Fordern? („Die Härte des logischen Muß")[14]

Eine solche Despotie ist eine – zumindest mögliche – Konsequenz der Searle'schen Intentionalitätstheorie. Wenn man sie umgehen könnte, würde man interessante Fragen nach dem Wechselverhältnis zwischen Intention und Erfüllung stellen können. Es würde auch Raum geschaffen für die Überlegung, ob nicht in den Erfüllungsbedingungen gar nicht enthaltene Objekte oder Sachverhalte den intentionalen Zustand erfüllen könnten. Doch eine solche Theorie kann in diesem Rahmen nur despotisch gefordert werden.

ANMERKUNGEN

1 Husserl, E., *Fünfte Logische Untersuchung*, hrsg., v. E. Ströcker (Hamburg 1975), S. 1.
2 op. cit., S. 37.
3 op. cit., S. 38.
4 s. op. cit., S. 102.
5 s. op. cit., S. 103.
6 Loon, J. van, „Rules and Commands", *Mind* 67 (1958), S. 524–521.
7 Hofstädter, A./Mc Kinsey, J. C. C., „On the logic of imperatives", *Philosophy of Science*, 6 (1939), S. 446–457.
8 Rescher, N., *The Logic of Commands* (London 1966).
9 Searle, J. R., Intentionality (Cambridge 1983).
10 op. cit., S. 1.
11 s. op. cit., S. 6.
12 op. cit., S. 11.
13 op. cit., S. 13.
14 *PU* § 437.

* * *

„KULTUR“ UND „VERHALTEN“ IN DER PHILOSOPHIE DES GEISTES

Peter Klein
Universität Hamburg

„Warum haben die Papua eine Kultur
und die Deutschen keine?“

(A. Loos)

Lebewesen sind energetisch offene Systeme, welche ihren spezifischen Systemcharakter dadurch sowohl entfalten wie erhalten, daß sie, nach einem Regelsystem mit dem Umfeld wechselwirkend, sich dessen Eigenschaften zunutze machen.

Bei Tieren nennen wir dieses System, unter Einschluß interner Bedingungen und Antriebe, ihr „Verhalten“. Es evolvierte *aufgrund* jener Wechselwirkung und sie optimierend, zugleich mit der betr. species, sie ebenso kennzeichnend wie deren morphologische Merkmale und mit ihnen eine funktionelle Einheit bildend, den „Bauplan“ (v. Uexküll).

Menschen als Lebewesen benötigen ebenfalls lebenssichernde Mechanismen. Diese sind jedoch, anders als beim Tier, reduziert zu Verhaltensdispositionen, die erst durch Lernen ergänzt manifestes Handeln ermöglichen und daher durch lokale Bedingungen und Traditionen modifiziert werden können. Dabei entsteht aber nicht beliebiges Verhalten, sondern es stabilisieren sich gut ausgrenzbare Einheiten zwischen species und Individuen, welche sich eben durch die unterschiedlichen Systeme ihrer Lebensregeln konstituieren, die „Kulturen“, definiert als Gesamtheit der Lebensäußerungen einer eben dadurch ausgrenzbaren Population einschließlich ihrer internen Antriebe und Vorstellungen (Mühlmann).

Die Bestimmungen von „Kultur“ des Menschen und „Verhalten“ der Tiere sind formal äquivalent. Ein möglicher Unterschied könnte ontologisch als Wesensunterschied von Mensch und Tier gedacht werden. Unter der Leitvorstellung einer evolutionären Deszendenz des Menschen vom Tier, mithin der Arbeitshypothese homologer Verwandtschaft ihrer Weltorientierungsleistungen, aber wäre diese Interpretation bloß formal. Unsere Analyse soll auf inhaltliche Unterschiede beider Begriffe gerichtet sein.

„Kultur“ schlechthin zu besitzen wird erst seit dem 18. Jh. allen rezenten menschlichen Gesellschaften zugeschrieben, weil ethnologische Forschung die Einsicht erbrachte, daß es in ihnen keine primitiven Vorstufen der Vernunft gibt, sondern nur unterschiedliche Realisierungen der einen Vernunft als potentia. Auf u.a. dieser Basis bestimmte auch schon das 18. Jh. diese Vernunft allgemein, am umfassendsten *Kant*.

Tierische Daseinsorientierung ist gekennzeichnet durch radikale Reduktion der in komplexen Situationen bedeutsam werdenden, d.h. das Verhalten beeinflussenden Information (Uexküll, Tinbergen): Wenige Anzeichen stehen mit genügender statistischer Geltung für ganze Situationen, setzen als „Schlüsselreize“ Wirkorgane in Tätigkeit, deren Funktionen genetisch fixiert, u.U. durch Lernen und interne „Stimmung“ modifiziert, ablaufen, deren ebenfalls bloß statistische Überlebensfunktion sich nur dem äußeren, systematisierenden Beobachter verrät, dem Tier aber beschränkt ist auf das Gefüge von Merk- und Wirkorganen, dessen innere Repräsentanz genau so weit reicht, wie das Tier in dies überlebensdienliche Gefüge eingebunden ist: Das Tier hat nicht „Welt“, sondern „Umwelt“ (vgl. weiterhin Klein).

Menschliche Vernunft dürfte somit dadurch zu kennzeichnen sein, dem Menschen „Welt“ zu vermitteln. Diese tritt aber nicht bloß als lebensdienlich perfektionierte, differenzierter wahrnehmende und bewirkende Repräsentanz von Reizen auf. Vielmehr ist es, mit Kant,