

**INTENTIONALITY, MODALITY
AND SUPERVENIENCE**

Hector-Neri Castañeda

Edited by

M. J. van den Hoven

and

G. J. C. Lokhorst

Erasmus University Rotterdam

Department of Philosophy

1990

This document was originally published as volume 13 of the
Rotterdamse Filosofische Studies (Rotterdam Philosophical Studies) of the
Faculteit der Wijsbegeerte (Department of Philosophy) at
Erasmus Universiteit Rotterdam (Erasmus University Rotterdam)
ISBN 90-70116-58-8

Conversion to pdf by
Gert-Jan C. Lokhorst <lokhorst@fwb.eur.nl>
on 15 July 1998

The page numbers in this document are different
from those in the printed book

CONTENTS

| | |
|--|----|
| FOREWORD | 7 |
| INTRODUCTION | 9 |
| QUINE'S EXPERIMENT WITH INTENSIONAL OBJECTS AND HIS EXISTENTIALIST QUANTIFIED MODAL LOGIC | 11 |
| Introduction | 11 |
| I. Quine's contribution to the semantics of quantified modal logic | 11 |
| I.1. Some of Quine's firmest convictions | 11 |
| I.2. The core of Quine's experiment | 12 |
| I.3. Quine's first investigation | 13 |
| I.4. Repudiation of material objects | 14 |
| I.5. Quine's major contributions: Individuals, sameness, and identity | 15 |
| I.6. Existence, identity, Russell's Robust Sense of Reality, and Pseudo-Quinean quantified modal logic | 17 |
| I.7. Individual constants | 20 |
| I.8. Definite descriptions | 21 |
| I.9. Essentialism and contingentism | 22 |
| II. Quine's 1961 assault on intensional objects | 23 |
| II.1. Referential opacity and convergent referential cumulation | 23 |

| | |
|--|----|
| II.2. Speaker's pragmatic "implication" of external by internal construal . . . | 26 |
| II.3. Referential opacity and quantification | 27 |
| II.4. Existentialist quantification vs. material objects | 28 |
| II.5. Quine's bomb against quantification over intensional objects | 30 |
| II.6. Is Quine's infinite regress vicious? Is it a regress? | 31 |
| II.7. Aristotelian essentialism | 33 |
| II.8. Pseudo-Quine's ontology and other sameness relations | 35 |
| III. Conclusion | 39 |
| Notes | 39 |
| Appendix | 41 |
| References | 43 |
| SUPERVENIENT PROPERTIES, EMERGENCE, AND | |
| THE HIERARCHY OF CONCRETE INDIVIDUALS | 45 |
| Introduction: Property supervenience, reduction, and emergence | 45 |
| 1. Supervenience | 47 |
| 2. The irreducible component of supervenient properties: Some questions | 49 |
| 3. The irreducible cores of supervenient properties: Equivalences vs. identity | 51 |
| 4. Individuals and supervenient reduction | 52 |
| 5. A look at the world | 55 |

| | |
|--|----|
| CONTENTS | 5 |
| 6. The hierarchy of individuals: The irreducibility of structure | 56 |
| 7. Mereological essentialism and world-object contingentism. | 59 |
| 8. Conclusion | 61 |
| Notes | 61 |
| INDEXICAL REFERENCE AND CAUSAL DIAGRAMS IN INTENTIONAL ACTION | 63 |
| Introduction: Four problems about intentional action | 64 |
| 1. Voluntary and intentional action | 65 |
| 2. The intentionality of intentional action | 68 |
| 3. Action, doing, characteristic state | 70 |
| 4. Goldman's analysis of intentional action: Praise and critique | 72 |
| 5. The main structure of intentional action: Data for any theory of intentional action | 77 |
| 6. First theoretical desideratum | 78 |
| 7. Causation and inference: conditional intentions | 78 |
| 8. Brand's fundamental question of action theory | 86 |
| 9. Intending is a generalization of sorts of believing | 86 |

| | |
|---|-----|
| 10. Volition | 89 |
| 11. The guiding element of an intention: The practical copula | 89 |
| 12. The mind-body relationship | 91 |
| 13. The internal causal diagrams | 92 |
| 14. Indexical reference in immediate volition | 94 |
| 15. The identification of items in one's perceptual spaces is demonstrative . . . | 96 |
| 16. The causal transferability of indexical volition | 97 |
| 17. External causal diagrams and the intersubstitutability of subordinate action | 98 |
| Notes | 100 |
| PHILOSOPHICAL PUBLICATIONS, 1986–1990 | 103 |
| Books | 103 |
| Main contributor in the following books in his honor | 103 |
| Essays | 103 |

FOREWORD

My semester (June–December, 1989) as the occupant of the Tinbergen chair at Erasmus Universiteit, assigned to the Faculteit der Wijsbegeerte, was very fruitful and very educational. Not only for me but also for my wife Rhina and my son Rodrigo. The experiences at Rotterdam were diverse and some of them very intense. We have made some profound friendships in Rotterdam. We will treasure our memories from Erasmus, Rotterdam, and The Netherlands.

The normal philosophical activity at Erasmus is very rich, very intense and very comprehensive. This is so both in teaching and in research. To illustrate, you find there experts on African philosophy, and on contemporary French philosophy, including the social philosophy of Emmanuel Mounier. To be sure, as a European university, there is a dominance of the history of philosophy, and, as one might have imagined, Erasmus has specialists working on the history of Dutch philosophy. And it is the great center for Spinoza studies, under Wim Klever, where absolutely nothing pertaining to Spinoza is considered unworthy of study.

The diversity of philosophical concerns at Erasmus made my visit a valuable opportunity for learning about other philosophical perspectives and ideologies. Experiences in which one sees one aspect of reality without its being molded by some of one's cherished presuppositions are always educational. Of course, it is impossible in six months to tap, let alone exhaust, all those fountains of knowledge.

Naturally, I was bound to have closer relationships with the so-called Analytic Philosophers of Erasmus, a small group of bright and devoted young philosophers: Jeroen van Rijen, Jeroen van den Hoven, Marjolein Degenaar, Gert-Jan Lokhorst, Frans Schaeffer, Hans Koeze, Hans Blom and George Berger. Our philosophical interaction was always pleasant and cordial, issue-delving and enriching. I was

fortunate in that they attended, and were active participants in, my lectures on human action.

The Faculty administration was very supportive of my work to the point of sponsoring an “intimate” European workshop on Objects, Reference, and Consciousness, as the culmination of my 1989 Tinbergen research appointment. The participants were from Holland, Italy, France, Switzerland, and Spain. Papers were read by George Berger, Eros Corazza, Jacques van Leeuwen, Paolo Leonardi, Gert-Jan Lokhorst, Kevin Mulligan, Francesco Orilia, Jesus Padillia Galvez, Francois Recanati, Marco Santambrogio. Jeroen van den Hoven helped secure the support of the International School for Philosophy at Leusden. Paul Wouters, the Director of the School, was splendid in his support of the workshop. Van den Hoven and Jeroen van Rijen were excellent both in their management of the preparations for the workshop and in their execution of their roles as moderators.

I am most grateful to all the members of the Erasmus Philosophy Faculty for their willingness to do what they could to make my visit to Rotterdam a joyous and valuable one. Dean Sperna Weiland and the faculty administration were very supportive. I am especially grateful to the “analytic” philosophers for their philosophical companionship and friendly courtesies. I am touched by the special care of Bert van Roermund, professor of philosophy of law at Tilburg, exercised in his monograph *Hector-Neri Castañeda: Een geschreven portret* (Tilburg, Tilburg University Press, 1989) both in reporting correctly biographical facts and in accurately displaying theoretical theses. Irma Beckers and Marlotte Schoorlemmer were superb in their assistance, as well as sources of information about Dutch culture.

I am grateful to Jeroen van den Hoven and Gert-Jan Lokhorst for the conception of this book, as a sort of public report of my work as the 1989 Tinbergen chair. The topics of these papers were treated in my lectures. I thank them also for the excellent way in which they have carried out the project – even spurring me to revise the papers and carefully editing them.

Hector-Neri Castañeda

INTRODUCTION

From June 1989 to December 1989, Hector-Neri Castañeda, The Mahlon Powell Professor of Philosophy at Indiana University, held the annual Tinbergen chair at Erasmus University Rotterdam. During this period, he was a visiting professor at the Department of Philosophy of this university.

When Castañeda arrived at the department, none of its members expected that his visit would have as great an impact as it eventually did. Everybody knew him by name, of course; but nobody suspected how much philosophical inquisitiveness, integrity, intelligence and intensity his name really stood for. In his courses, Castañeda weekly gave us an example of how the true philosopher should proceed. Castañeda's philosophical thought had no discernible horizon; in the spirit of the motto of the journal he has founded, *Noûs – nihil philosophicum a nobis alienum putamus* – no philosophical subject was too far out for him. On the other hand, the comprehensiveness of his interests did not stand in the way of profundity: there was not a single topic which he did not attack with equal acumen, rigour and originality. In short, Castañeda demonstrated that violation of the precept *multum non multa* does not necessarily have bad effects. His exemplary behaviour will forever be with us, and probably provide us with a bad conscience for the rest of our lives.

This collection of previously unpublished papers is intended as a *souvenir* of Castañeda's visit to Rotterdam. The papers show both the wide scope and the depth of his thought. They range from topics as abstract as logic and ontology, to such practically relevant areas as action theory, ethics, and the physiological basis of human behavior. Castañeda lectured on all these subjects during his stay. We hope the present collection may be as inspiring to the reader as the author's own presence has been to us.

The first essay is concerned with Quine's ideas about intensional objects and modal logic. As is well known, Quine has always been an opponent to modal logic. Nevertheless, in the process of reducing it *ad absurdum*, he implicitly made some interesting suggestions about it himself. These insights are extracted by Castañeda and shown to be anticipations of some of the theses of his own guise theory.

The second paper deals with the currently much-discussed concept of supervenience. Castañeda presents a foundational study of some basic ontological and epistemological issues involved in the supervenience thesis – the thesis that although it is not true that everything is reducible to physics, physical reality nevertheless *determines* the whole of reality. Castañeda argues that the supervenience thesis gives an impoverished view of reality: our ordinary world, the world we find ourselves living in, is structured by us in ways which may be diametrically opposite to the way in which physics organises reality – even though the world may be wholly physicalist at bottom.

The third article contains a description of the causal mechanisms which are involved in intentional action. Castañeda shows that his conception of human agents as “indexically volitional” agents facilitates an account of the internal processes preceding voluntary action. In this respect, the paper is a valuable addition to the more comprehensive treatment in his *Thinking and Doing* (Dordrecht, Reidel, 1975).

Jeroen van den Hoven

Gert-Jan Lokhorst

CHAPTER 1

QUINE'S EXPERIMENT WITH INTENSIONAL OBJECTS AND HIS EXISTENTIALIST QUANTIFIED MODAL LOGIC

Introduction

Willard Van Orman Quine has always had firm philosophical convictions. Nevertheless, he has been attentive to alternatives, and has experimented with views for which he has felt little sympathy, even strong antipathy. Here are some philosophical reflections on his philosophical experiment with so-called intensional objects and their identity, which he performed in connection with his deeply-rooted animadversions against modal logic. The major experiment, conceived as a sort of *reductio*, is described in Quine 1947. Interestingly enough, in the process Quine hints at an existentialist quantified modal logic. Unfortunately, he did not bring the experiment to its completion; in fact he disowned it in Quine 1961. In this work he explodes a logical bomb in the heart of modal logic. These are just two episodes in the history of Quine's relationships to modal logic. The rest of this history is, however, beyond the present study.

I. Quine's contribution to the semantics of quantified modal logic

I.1. Some of Quine's firmest convictions

Quine has held most firmly that the world is composed of physical objects, ultimately the physical individuals posited by physics. Such and classes thereof are the preferred values of his quantifiers. Another firm conviction of his has been

Bertrand Russell's *robust sense of reality*, namely, that quantification and predication are the loci where language links to reality. Fittingly, he interpreted the quantifiers objectively. This interpretation underlies several of his other convictions. On the one hand, it supports his conviction that the values of the quantifiers are existents; hence, everything that exists, and the particular quantifier, which expresses membership in the domain of discourse, is *the* existential quantifier. Merely possible and impossible objects are by definition not real, and cannot be the values of quantification. On the other hand, at least during the period under consideration he relentlessly opposed interpreting the quantifiers substitutionally, even when the substituends denote, that is, pick up real values. These profound convictions undergirded his firm conviction that modal logic was built on a confusion of use with mention.

1.2. The core of Quine's experiment

Notwithstanding his firmest convictions, Quine experimented with the idea that the deep-seated confusion at the root of modal logic could be construed as the introduction of a domain of thinner individuals, carved out of the different ways of mentioning physical objects. In brief, he sets out to investigate the following experimental assumption:

(EA) Let the quantifiers range over such thinner individuals as would be the hypostases of the different ways of referring to physical objects.

Quine does not put his experiment exactly in these terms; he couches it, rather, as a study on the substitution interpretation of the quantifiers. He, thus, signals the connection between his deep convictions about objectual quantification and about modal logic. He considers the following truth criterion (Quine 1947, 46):

(TC) An existential quantification holds if there is a constant whose substitution for the variable of quantification would render the matrix true.

He points out that this “is at best a partial criterion (both in modal and in non-modal logic), because of unnameable objects”; then he investigates “how it fares” as a sufficient condition.

Clearly, if the ways of referring to objects as individuals involve making use of individual constants, then (EA) is tantamount to (TC). We discuss individual constants below.

1.3. Quine's first investigation

In an experimental vein Quine explores two basic tenets of modal logic as these apply to astronomy:

- (1) The Morning Star is the same as the Evening Star;
- (2) Not necessarily (the Morning Star is the same as the Evening Star);
- (3) Necessarily (the Morning Star is the same as the Morning Star).

In a move with a hint of a *reductio ad absurdum*, Quine proposes to take the difference between (1) and (2) seriously. He calls *congruence* the sameness, which “according to empirical evidence” (Quine 1947, 47), makes (1) true. He offers the symbol '*C*' as the canonical representation of (empirical) congruence. Then he interprets the locution 'is the same as' in (3) as *also* expressing congruence. He symbolizes (1)–(3), where '*N*' stands for C.I. Lewis's necessity box, which Quine uses, thus:

(1.Q) $MS C ES$.

(2.Q) $\sim N(MS C ES)$.

(3.Q) $N(MS C MS)$.

By parity with (3.Q) we have:

(4.Q) $N(ES \ C \ ES)$.

Then Quine, conjoining and applying criterion (TC) to the “names” abbreviated as 'MS' and 'ES', derives:

(5.Q) $(\exists x)(x \ C \ ES \ \& \ N(x \ C \ MS))$ {from (1.Q) and (3.Q)}.

(6.Q) $(\exists x)(x \ C \ ES \ \& \ \sim N(x \ C \ MS))$ {from (1.Q) and (2.Q), and the symmetry of C}.

By Leibniz's identity of indiscernibles, clearly *ES* and *MS* are distinct objects.

Obviously, the argument is generalizable to any “name” or way of referring to Venus, including the name 'Venus' itself. Quine concludes (*Op. cit.*, 47, the title and the italics are mine):

QUINE'S METATHEOREM FOR MODAL SEMANTICS

Thus it is that the contemplated version of quantified modal logic is committed to an ontology which *repudiates material objects* (such as the Evening Star properly so-called) and leaves only multiplicities of distinct objects (perhaps the Evening-Star-concept, the Morning-Star-concept, etc.) in their place.

This is the reduction to absurdity hinted at. Quine sees that this consequence may not seem so absurd to modal philosophers with different convictions from his own. However, he counts on some modal philosophers finding the “repudiation of material objects (or, indeed, of classes) uncongenial.”

1.4. Repudiation of material objects

As far as the 1947 Quine is concerned, the “repudiation of material objects by quantified modal logic” was the immediate consequence of the use/mention confusion inherent in modal logic. For him that “repudiation” was sufficient to

establish the irredeemable philosophical bankruptcy of modal logic. He did not consider the reply that material objects are not repudiated, but only *analyzed* as equivalent *C*-systems of finer-grained individuals. Although the 1947 Quine does not say so, he is writing from his firm reductionist conviction that “to analyze is to analyze away.” This slogan, like all robust reductionisms, evinces the view that somehow the complex does not exist. In any case Quine's major 1947 objection to modal logic seems to depend squarely on his views of analysis. Non-reductionists were bound to remain unpersuaded.

Quine was, however, from his own perspective, justified in accepting the argument as a kind of *reductio*. First of all, the above mentioned “analysis” conflicted with some of his firmest-grounded convictions. After all serious philosophers philosophize in the first person for the first person. Second, given the indeterminacy of theories – understanding of which we owe to a large extent to Quine himself – , Quine was doxastically justified in taking material objects, or the ultimate objects of physics, and classes thereof, as his primitive ontological category.

Modal logicians and philosophers may, conversely, also have a right to their views. (This echoes Carnap's Principle of Tolerance.) They must, however, develop them consistently, and make them more and more comprehensive. In this spirit they would, and should, welcome Quine's experiment as revealing part of the structure of the world to which their modal views commit them. (Perhaps, as Quine intimated, that revelation might be effective propaganda against such views.)

1.5. Quine's major contributions: Individuals, sameness, and identity

Quine's experiment goes directly to the heart of the ontology of modal logic. His proposal to consider a special relation of contingent sameness should be welcomed by modal philosophers. Quine's congruence is precisely the empirical sameness which allegedly a Babylonian astronomer discovered to hold between the morning star and the evening star. This sameness should, most emphatically, be distinguished from identity, i.e., self-identity, which is necessary – and *a priori* because it is so trivial. Of course, both samenesses could be defined in terms of a generic sameness and the modalities.

In brief, in 1947 Quine seems to have made two major contributions to modal logic, by establishing the following:

TWO CRUCIAL FEATURES OF THE NATURAL SEMANTICS OF
QUANTIFIED MODAL LOGIC

- (I) *Fine-grained individuals*. A genuine quantified modal logic, i.e., one that does not collapse into standard first-order predicate logic, has as values of its quantifiers fine-grained individuals, finer-grained than material objects. Let us for convenience call these fine-grained individuals *individual guises*, in general, and *material guises* those composing material objects.
- (II) *Two sameness relations*. Such a logic requires of course a corresponding universal necessary identity relation that, by applying to all such fine-grained objects, becomes itself fine-grained; however, it also requires a contingent *coarse sameness* relation, which gathers together the fine-grained individuals into coarse-grained material objects. Whereas the necessary identity should yield an unrestricted principle of substitution in all contexts of the modal language in question, the coarse sameness should allow just an extensional principle of substitution valid for non-modal contexts only.

Here we can see the quiet operation, in the background, of Quine's principle: *No entity without identity*.

Modal philosophers would, however, demur at Quine's interpretation of (3) as (3.Q). They would likely hold that the very existence of stars, let alone the morning star, is a contingent matter. Hence, what (3) says is:

(3.N) $N(MS = MS)$.

Quine's *C* is a contingent equivalence relation among existents. Hence, intended as (3.Q) sentence (3) is false.

I.6. Existence, identity, Russell's Robust Sense of Reality, and Pseudo-Quinean quantified modal logic

Quine was, however, from his own view of quantification, justified in building his argument on (3.Q). His view of quantification requires that the values of the variables of quantification be real, existing individuals. His argument seems to presuppose the following tacit *existentialist postulate*:

$$(Q.=C) x = y \rightarrow xCy.$$

Typically in modal logic with identity:

$$(N=) x = y \rightarrow N(x = y).$$

$$(N\rightarrow) N(p \rightarrow q) \rightarrow (Np \rightarrow Nq).$$

Given that Quine was not particularly interested in fostering modal logic, the implicit formula (Q.=C) should perhaps be called *Pseudo-Quine's*, or *C-Quine's*, postulate. It has the effect of making the objects in the domain of quantification “necessary” existents. (Q.=C), central to Quine's cryptic existentialist quantified modal logic, pivots on Russell's *Robust Sense of Reality*. This underlies Quine's philosophy of logic. It includes not only the tenet that the domain of quantification is a domain of existents, thus, representing existence with variables binding to quantifiers, and equating *a exists* with $(\exists x)(x = a)$. It also includes the Parmenidean thesis that non-existents have no properties, which grounds Russell's equating *a exists* with *a is F* for some positive nonlogical predicate, for instance, an atomic or primitive predicate. Full specificity is gained by selecting any universally applicable predicate, for example, the identity predicate; thus Russell equates *a exists* with $a = a$. We refer to these equations as *Russell's conflations*.¹

Clearly, (Q.=C) is not ontologically satisfactory. Physical objects exist contingently. At most, perhaps, it is necessary that there exist some physical object or other, but not any particular one. *Pseudo-Quine* might, perhaps, have defended his

existentialist modal logic as epistemologically correct. Ontology being a matter of linguistico-philosophical policies, (Q.=C) simply represents his decision to deal with existents only. Had he noticed it, Quine might have regarded this ambiguity of (Q.=C) as another objection against modal logic.

Patently, a modal philosopher-logician who recognizes the contingency of existence will claim that the new perplexity about (Q.=C) is solved by fastening to the latter's falsehood, which itself establishes this:

THIRD CRUCIAL FEATURE OF THE NATURAL SEMANTICS OF QUANTIFIED MODAL LOGIC

(III) *Quantification over nonexistent individuals.* Such a logic, to be really useful, must have its most basic quantifiers ranging over larger domains of values than domains of existents. The existential sub-quantification ranges over a proper subdomain, and existence becomes a differentiating property. This property may be conceived as Quine's core of *self-congruence* once postulate (Q.=C) is jettisoned. Hence, with non-Quinean variables of quantification (say, 'g', 'g_i', ...) taking possibles (even impossibles) as values, existence can be defined in *possibilist* modal logic extended by the adjunction of Quine's sign 'C' as self-congruence, thus:

(Ex.C) *g exists: E!g = Def. gCg.*

Other formulas that undoubtedly should be theorems about existence or congruence are the following, where 'a', 'b', and 'c' are singular terms or individual variables:

(C.Sym) $aCb \rightarrow bCa.$

(C.Ref) $aCb \rightarrow aCa.$

(C.Tr) $aCb \rightarrow (bCc \rightarrow aCc).$

The possibilist modal logician rejects the possibilist generalization of *Pseudo-Quine's* postulate (Q.=C), namely: $g_1 = g_2 \rightarrow g_1 C g_2$.

Patently, $g_1 C g_2$ may be false, even if $g_1 = g_2$, by g_1 and g_2 failing to exist, indeed they might even be impossible.

A brief note on an application of modal logic. The possibilist modal philosopher of science may deem quantification over merely possibles (impossibles, too?) to be a distinctive central mechanism of a fully-fledged modal logic. On his view, that mechanism makes modal logic a useful tool in considering how to enrich the domains of scientific theories, while remaining open to the possibility that some of the posited entities may turn out not to exist – indeed some may be impossible.

In any case, the distinction between domains of possibles and domains of existents is orthogonal to the distinction Quine argued for, recorded in (I) above, between material objects and material guises. Quantification over possibles, whether fine-grained or coarse-grained, threatens to tear the 1947 Quine's Russellian robust sense of reality much more deeply than quantification over (existing) fine-grained individuals. This horror of possibilia is precisely what the existentialist postulate (Q.=C) evinces – under the provision that there are *no* more general variables than Quine's 'x', 'y', etc., which are restricted to existents. Thus whereas the fine-grained material guise The Morning Star is in (*Pseudo-*)Quine's domain of quantification, the equally fine-grained guise The Noon Star is not.

The 1947 Quine might have been willing to experiment with just one exception to Russell's conflation, to wit, letting nondenoting terms flank true identities. The reflexivity of identity is truly universal, and '=' is a genuine logical sign. This matches the fact that nondenoting terms enter in logically true sentences. *Pseudo-Quine* may, thus, accept $a = a$ as true, but not fa for atomic f , when ' a ' does not denote. Objectually speaking, nonexistents can have only logical properties. For example, if f is an atomic property and a does not exist, and fa is false, a satisfies the property $fx \vee \sim fx$. Existence is also equated with having at least one contingent atomic property.

1.7. Individual constants

Quine, in his substitution criterion of truth (TC) above, speaks of (individual) constants, and refers to them as “names” when he argues against the necessity of (TC): some objects are unnameable. He uses the names 'Morning Star', 'Evening Star', and 'Venus', rather than the definite descriptions 'the morning star,' 'the evening star', or 'the thing that Venusizes'. For Quine, however, proper names and individual constants are dispensable: they are analyzable in terms of definite descriptions. This raises the question whether, for the sake of his experiment with intensional objects, the 1947 Quine considered, or would have considered, definite descriptions as individual constants. This, to be sure, would have conflicted with his strong conviction that definite descriptive phrases are analyzable in Russell's way, in terms of quantifiers and identity, for example as follows:

(R.Def.Des) $Z\{Y(\iota xFx)\} = \text{Def. } Z\{(\exists x)(Fx \ \& \ (y)(Fy \rightarrow y = x) \ \& \ Yx)\}$, where
'{...}' is the scope of ' $\iota x(Fx)$ '.

Evidently, then, the experimental *Pseudo*-Quine has to differentiate himself from Quine more and more. He must drop either Quine's dispensability of names, or the Russellian analysis of definite descriptions. Otherwise there are no constants for the modal logician to apply criterion (TC).

The dispensability of proper names, or arbitrary individual constants in favor of definite descriptions seems to be sound. Doubtless, proper names are not in general, or semantically, reducible to uniquely applicable predicates, as Quine sometimes has urged. Notwithstanding, they are predicates, indeed they are in principle applicable to an indefinite number of objects; in certain *contexts of use*, however, they gain uniqueness of application.² Of the many Quines in the world, the present context has determined the one and only Quine the preceding discussion is about. In fact, the discussion so far has been about one Quine slice: the 1947 Quine. This of course is not quite correct according to some modal philosophers. We have, according to them, been discussing an individual guise: the 1947 philosopher who was a Quine, experimented with the semantics of quantified modal logic, and wrote Quine 1947.

Let then the 1947 *Pseudo-Quine* allow definite descriptions as primitive individual constants. Hence, in the modal logic under consideration the constants for criterion (TC) are of the form $\iota x(Zx)$, for any well-formed formula Zx . These constants may of course fail to denote. Yet, by the previous exception to Russell's conflation of existence with the predication of identity, we have $N(\iota x(Zx) = \iota x(Zx))$ as a theorem. On the other hand, $\iota x(Zx) \text{ C } \iota x(Zx)$ is equivalent to $E!\iota x(Zx)$, representing $\iota x(Zx)$ *exists*, and neither of them is necessarily true, much less theorems of modal logic.

1.8. Definite descriptions

Recall that Quine's experiment pivots on his meta-theorem that, given the substitution criterion (TC), in quantified modal logic distinct individual constants or names must denote different fine-grained values of the quantifiers. This is self-evident if the modal individual constants are definite descriptions: they have a palpable difference in the different predicates forming them. For example, it is a wholly empirical matter whether the F , which we may suppose to exist, is the same as the $F \& G$, for some contingent property G . Clearly, this sameness cannot be self-identity; if it holds it must be congruence.

In short, *Pseudo-Quine's* existentialist quantified modal logic is a sort of modal “free logic”: it contains nondenoting terms, but quantification is over existents only. It has as theorems formulas like the following:

(D. \sim =) $\sim(\iota x(Zx) = \iota x(Yx))$, whenever Zx and Yx are different matrices.

(D.UI) $(x)(Fx \rightarrow (E!\iota x(Zx) \rightarrow F\iota x(Zx)))$.

Clearly, (D. \sim =) is the representation of Quine's above meta-theorem. (D.UI) is the “free logic” principle.

Pseudo-Quine abandons Russell's analysis of ' $\iota x(Zx)$ ' in order to provide his modal logic with the needed individual constants. However, weaker counterparts of

it should hold. For example the following, which captures the existential assumption of Russell's analysis, should be forthcoming as a theorem:

$$(M.DD.Ex) E! \iota x(Zx) \leftrightarrow (\exists x)(Zx \ \& \ (x = \iota x(Zx)) \ \& \ (y)(Zy \rightarrow xCy)).$$

Likewise, the Parmenidean assumption should be captured by a theorem like:

$$(M.Pred) \sim N(x)(Zx) \rightarrow (E! \iota x(Zx) \leftrightarrow Z \iota x(Zx)).$$

1.9. Essentialism and contingentism

As we have seen, in *Pseudo-Quine's* existentialist modal logic, an atomic predication-matrix fx represents existential predication. Hence all individuals that make it true both exist and satisfy it contingently; that is, $\sim N(fa)$ even if fa , which implies $E!a$. As we have discussed, that holds even when $a = \iota x(fx)$. The only essential properties that existents – as well as non-existents – have are logical properties.

Clearly, the situation is not wholly satisfactory. Sometimes some philosophers want to say that (4) below is necessarily true:

(4) The present queen of England is a queen.

This necessity *Pseudo-Quine* does not allow. He is of course just a modal heir of Quine and Russell. Interestingly enough, other philosophers, for example followers of Kant, are anxious to claim not only that (4) is necessarily true, but also that it is so regardless of the existence of the present queen of England. For them “The 1989 king of the State Arizona is a king” is also necessarily true.

A modal possibilist logician provides logical space for individuals possessing contingent properties essentially. He gives up the existentialist view of predication with his possibilist quantifiers. In particular, a possibilist extension of *Pseudo-Quine's* existentialist logic has a guise of the form $\iota g(fg)$ *possesses essentially the property f-ness*. That is, he has $N(f \iota g(fg))$. He might define *Pseudo-Quine's* existential predication as follows:

(Ex.Pred) $fx = \text{Def. } (\exists g)(gCx \ \& \ fg)$.

Using the same kind of general guise variable, the possibilist could perhaps define the existentialist predication as the existential modalization of his general predication: Cfg as short for $(\exists h)(gCh \ \& \ fh)$.

In this possibilist quantified modal logic all guises have-essentially their internal properties, but the existing ones also have-contingently all their properties.³

II. Quine's 1961 assault on intensional objects

II.1. Referential opacity and convergent referential cumulation

Quine 1961 (130–159) contains a limpid exposition of Quine's firmest convictions about reference, quantification, objects, and modal logic. It summarizes, only too briefly, his 1947 experiment with intensional objects, and charges against modal logic with new artillery. His discussion is broader: he treats alethic modal logic and the psychological modalities in a unitary way. Hereafter we will for the most part concentrate on the psychological modalities.

Quine nails down forever the crucial fact that both types of modal construction have, characteristically, positions of reference that are, as he calls them, *referentially opaque*. He makes a very profound point about the intrinsic semantico-syntactic nature of the alethic and psychological modalities. He explains referential opacity as failure of valid inter-substitution of identicals, as failure of valid existential generalization, and as failure of inter-substitution of equivalents. These criteria are, however, not equivalent. (See Sharvy 1972.) Here I explain Quine's deep point about the semantics of psychological modalities in terms of the *actual use* of psychological sentences to make statements, whether these be true or false. My explanation provides a *pragmatic grounding* of Quine's logical claims for the psychological modalities.

Let's ponder the following example:

(5) Columbus believed that Queen Isabella's lover was French.

We focus on the term 'Queen Isabella's lover'. Sentence (5) can be used to make two different kinds of statement, depending on how the term is construed. In part this difference in construal is a difference in logical scope: in *internal construal* the term 'Queen Isabella's lover' as such lies wholly in the scope of 'Columbus believes that'; in *external construal* the term as such lies without the scope of 'Columbus believed that'. However, the term plays in (5) other roles. These other roles bring in external aspects to the internal construal, and internal aspects to the external construal. In short, each construal is a manifold of aspects. This is why we cannot identify occurring in internal construal with occurring *de dicto* – to use the medieval expression Roderick Chisholm has used to capture internality. Another reason is that the *de dicto/de re* contrast has, in becoming standard, developed ambiguities. Quine's referential opacity, I submit, is a consequence of internal construal. To see this let's return to the pragmatic exegesis of (5).

(A) *Internal construal*. With the term so construed, sentence (5) can be uttered to make a statement through which the speaker *attributes* to Columbus possession of the expression 'Queen Isabella's lover' or its counterparts in some unspecified languages. On this interpretation if Columbus thought out loud what according to (5) he believed, he proffered a sentence in one of the languages he spoke, which sentence would translate as "Queen Isabella's lover is French." The clause in internal construal represents Columbus's *dictum*, what he said, would have said. The internal construal of a term has, however, its characteristic external strands, which provide it with a peculiar *cumulative referential role*, to wit:

(a) Certainly to the person she is speaking of the speaker of (5) attributes the possession of a definite mechanism of referring, which is the term itself or a counterpart in some language available to that person;

(b) But she herself uses the internal expression as a means of depicting within her own experience and language what the person she is speaking of refers to, or could refer to;

(c) Thus, the speaker is herself thinking what she attributes;

(d) Furthermore, terms in internal construal depict the speaker's assumption of a referential convergence between herself and the thinker she speaks of.

To illustrate, the speaker of (5) with internal construal depicts the speaker's supposed agreement with Columbus in referring, or at least in purporting to refer, to one and the same entity as Queen Isabella's lover.

Clauses composed of expressions wholly in internal construal are, semantically, like verbal *pictures* the speaker draws of the thought contents of the persons she is speaking of. They are pictures within her own language: she understands them: they are *her* own representations of others' contents. The cumulateness of internal construal is principally representational cumulateness.

(B) *External construal*. A speaker of (5) with the term construed externally makes a statement in which she *alone* calls something "Queen Isabella's lover"; she refrains from attributing to Columbus any particular way of referring to that something, indeed even the idea that Isabella had a lover. The externality of the external construal of a term in modal sentences is nothing else than the term itself carrying simply the *speaker's reference* without cumulation of reference and attribution. Thus, a speaker of (5) with external construal merely represents Columbus as having said, in case he, again, thought out loud what he is said to have believed, something of the form: *Alpha is French*, where 'alpha' is a stand-in for some unspecified name, description, or indicator, which Columbus could have used to refer to what the speaker calls Isabella's lover. Thus, on the external construal of the term, (5) has the following preliminary analysis:

(5.B) Speaker: Queen Isabella's lover

\\

Columbus believed that *alpha* was French.

What goes on when a person asserts (5) includes the following:

(a) The speaker of (5) makes her own *unshared* reference to what she calls "Queen Isabella's lover";

(b) she signals that from her perspective her Queen Isabella's lover is the *same* as something Columbus (could have) referred to in some so far unknown ways, which for convenience we represent with the word 'alpha';

(c) she submits that the unspecified mechanism 'alpha' would have played in Columbus's utterances is the same as the logical role 'Queen Isabella's lover' plays in the speaker's sentence "Queen Isabella's lover was French";

(d) this logical role is played at the internal position marked by *alpha* in (5.B), which 'Queen Isabella's lover' occupies.

Clearly, that internal position is occupied by items in internal construal as well as by items in external construal. That position represents the speaker's *assumed* convergence of her own reference and the reference, specified or not, she attributes to Columbus. Thus, external construal presupposes internal construal.

Internal positions are characteristic and peculiar features of psychological constructions. They are the necessary pivots of the depicting role of clauses subordinate to psychological verbs. These clauses do not really depict tokened or tokenable sentences; they depict thought content or mental representation. Let's call truth-valued thought contents *propositions*. The characteristic and peculiar function of subordinate clauses with all its expressions in internal construal is to depict the propositions that are the contents of mental states attributed to others. The chief function of clauses in indirect speech is to reveal propositions, or propositional components, or at least propositional structure. Internal construal is, thus, *propositionally transparent*. Because of the referential cumulateness of expressions in internal construal, they cannot express the speaker's unshared reference. Hence, they must be, in Quine's Russellian phrase, referentially opaque.

II.2. Speaker's pragmatic "implication" of external by internal construal

As we have seen, there is an important and intimate connection between internal and external construal. Because of its cumulateness, an internal construal includes the speaker's own (purported) reference blended with her attribution of reference. She may take her own reference out of that internal blend, if she wishes to do so. The requirement is obvious: she must feel *comfortable* with the item itself which is the target of the referential cumulation. This comfortableness involves more than merely sharing the same ontology; it includes sharing somehow a view of the item in question. For example, the speaker of (5) may also believe that Isabella had a lover.

Then she can withdraw her own reference from the internal referential bundle at 'Queen Isabella's lover' in (5) and claim:

(5.A-SW) Columbus believed that Queen Isabella's lover was French. But he [or Queen Isabella's lover] was actually a Sardinian at the Spanish Court in Naples.

Here is a sort of *pragmatic implication* available to the speaker. That is, a psychological statement with a subordinate clause with terms in internal construal pragmatically "implies" *in the speaker's speech* the statement with external construal, provided that the speaker satisfies the presupposition of shared belief or experience.

II.3. Referential opacity and quantification

Quine defines referential opacity in logical terms as the failure of substitutions and existential generalization at opaque terms or positions. In the case of the psychological modalities such failures can be readily explained by conjecturing that Quine has implicitly adopted the speaker's point of view. To begin with, the terms in external construal carry the speaker's reference only; she uses them to place the objects they denote in her world: these objects are, therefore, in her own domain of discourse and she is free to say so by means of the particular or existential quantifier. Consider, for example, the speaker of (5) with external construal (B): she, *not* Columbus, refers to Queen Isabella's lover. The speaker's utterance represents the speaker, not Columbus, as believing that Isabella had a lover. She can infer from (5) on interpretation (B) that:

(6) $(\exists x)$ (Columbus believed that x was French).

All this Quine accepts; it is not what he calls 'quantifying in'.

On the other hand, the internal construal (A) of (5) attributes to Columbus the belief that Isabella had a lover. Now trouble appears. Suppose that the speaker does

not herself believe that Isabella had a lover. Then she *cannot* quantify existentially on that term and honestly assert (6) above, or for that matter (7) below:

(7) $(\exists x)(\text{Same}(x, \text{Queen Isabella's lover}) \ \& \ \text{Columbus believed that } x \text{ was French})$.

The 1961 Quine ends his own discussion of examples with the following convincing diagnosis:

The root of the trouble was the referential opacity of modal contexts. But referential opacity depends in part on the ontology accepted ... (Quine 1961, 152).

Quine's specific conclusion is that quantification over coarse-grained material objects is the obstacle to making sense of quantified modal logic, that is, of quantification over internally construed terms or vacant (free) internal positions. The problem is especially serious for the case of psychological statements. These formulate attributions of mental state, and, as we have seen, their peculiar and common goal is to attain some degree of propositional transparency. This requires the occurrence of expressions, at least those signaling propositional structure, to occur with internal construal. This generates referential opacity. Thus, referential transparency is an *aberrant* phenomenon that occurs within a background of propositional transparency and referential opacity.

II.4. Existentialist quantification vs. material objects

I am not sure, however, that the difficulty lies solely in the domain of quantification being composed of material objects. Suppose that we quantify over a domain of fine-grained individual guises, but follow Quine in employing existentialist quantifiers. For example, suppose that 'Queen Isabella' and 'Queen Isabella's lover' denote guises, or Fregean senses, or Carnap's individual concepts. (These are neither exactly identical nor guises of each other.) If the particular quantifier has the existentialist

interpretation, then non-existing guises (concepts, or senses) would be outside the domain of quantification. Hence, if Queen Isabella's lover does not exist, neither (6) nor (7) follows from (5), whether interpreted with the term in internal or in external construal.

Psychological-modal quantification over terms construed internally cannot run smoothly unless its domain is extended *both* to range over fine-grained individual guises *and* to range also over nonexisting guises. An ontology that is both fine-grained and possibilist seems to be required for the full benefit of modal psychological logic. This is the psychological counterpart of the THIRD CRUCIAL FEATURE of alethic-modal semantics recorded above in section I.6.

To illustrate let us return to (5) with its terms in internal construal. On possibilist and fine-grained quantification, regardless of whether the speaker believes, or fails to believe, that Queen Isabella had a lover, (5) implies:

(6.G) $(\exists g)(\text{Columbus believed that } g \text{ was French})$.

(7.G) $(\exists g)(g = \text{Queen Isabella's lover} \ \& \ \text{Columbus believed that } g \text{ was French})$.

A modal possibilist speaker can, it seems, move with a clear conscience from asserting (5) understood with internal construal – which is referentially cumulative – , to asserting it with external construal, to making other statements about Queen Isabella's lover, and then proceed to deny that lover's existence, as follows:

Columbus believed that Queen Isabella's lover was French. So Columbus thought of Queen Isabella's lover as French. Admiral Gonzalez thought that he was Italian. He has been praised and reviled. Everybody wanted to know him. But that couldn't be. Queen Isabella's lover never existed. Hence, her French lover Columbus thought of never existed either. Well, here is something talked about so much that does not exist.

To be sure, this modal speaker cannot consistently use Quine's existential variables: he should not be asserting $(\exists x)(\sim E!x)$. He must, rather, be using his own possibilist variables and asserting:

$$(8) (\exists g)(\sim E!g).$$

Aiding himself to Quine's existential variables, our modal speaker can formalize his statement (8) as: $(\exists g)(\sim(\exists x)(x = g))$. In terms of Quine's 1947 existential congruence, (8) is $(\exists g)(\sim gCg)$. Of course, neither the 1947 Quine nor *Pseudo-Quine* endorses the use of such general variables.

II.5. Quine's bomb against quantification over intensional objects

In 1961 Quine has, however, lost all patience with modal logic. He is no longer interested in exploring – even tongue in cheek – the inner structure of modal logic. He does not even mention his early experimental congruence, *C*. The reason seems to be twofold. First, his argument for the fine-grained ontology in quantified modal logic is more general, no longer dependent on congruence. Second, he has found an impressive explosive argument which he thinks reduces quantified modal logic to shambles in one simple burst. The argument constructs an infinite regress for quantified modal logic. Here is Quine's assault in full detail (Quine 1961, 152–153):

Actually, even granted these dubious entities, we can quickly see that the expedient of limiting the values of variables to them is after all a mistaken one. It does not relieve the original difficulty over quantifying into modal contexts [i.e., into positions with internal construal]; on the contrary, examples quite as disturbing as the old ones can be adduced within the realm of intensional objects. For, where *A* is any intensional object, say an attribute, and '*p*' stands for an arbitrary true sentence, clearly

$$(9) A = \iota x [p \ \& \ (x = A)].$$

Yet, if the true sentence represented by ' p ' is not analytic, then neither is (9), and its sides are no more interchangeable in modal contexts than are 'Evening Star' and 'Morning Star', or '9' and 'the number of planets'.

Or, to state the point without recourse to singular terms, it is that the requirement lately italicized – “any two conditions uniquely determining x are analytically equivalent” – is not assured merely by taking x as an intensional object. For, think of ' Fx ' as any condition uniquely determining x , and think of ' p ' as any nonanalytic truth. Then ' p & Fx ' uniquely determines x but is not nonanalytically equivalent to ' Fx ', even though x be an intensional object.

II.6. Is Quine's infinite regress vicious? Is it a regress?

Quine rightly claims that his (9) leads to contradictions exactly parallel to those of the Morning Star and Evening Star in Part I above. Now, the 1947 *Pseudo-Quine* eliminated such contradictions in three steps. He distinguished, first, between necessary identity (=) and contingent congruence (C); second, he interpreted the argument leading to the contradiction as a *reductio* of $MS = ES$; third, he urged us to take as true $MS C ES$ instead.

The 1947 *Pseudo-Quine* could continue his experiment and treat the 1961 Quine's paradox hinging on (9) in the same way. He would take Quine's argument to show that (9) is false, and would propose that the truth confused with (9) is the C -formula at the end of:

$$(9.P-Q) p \rightarrow (E!A \leftrightarrow A C \iota x[p \ \& \ (x = A)]).$$

This interpretation lies within *Pseudo-Quine's* existentialist view of quantification.

Concerning Quine's argument in terms of variables, *Pseudo-Quine* would reply that sentential matrices have two functions in his modal logic. On the one hand they determine individual constants, which are primitive definite descriptions. On the other hand, they are truth-valued schemata. Hence, first, the matrices ' Fx ' and ' p & Fx ' uniquely determine, respectively, the individual constants, ' $\iota x(Fx)$ ' and ' $\iota x(p \ \& \ Fx)$ '

Fx '. As the matrices are not “analytically” equivalent, the ensuing constants denote, or, rather, purport to denote, two *different* fine-grained objects. These are structurally and epistemically different: they are distinct thinkables, lying within, or outside in the periphery of, the existentialist domain of quantification. They are, however, also semantically distinct. The matrix ' $N(y = \iota x(Fx))$ ' is true of $\iota x(Fx)$, but not of $\iota x(Fx \ \& \ p)$. Second, as contingent truth-value schemata, the matrices ' Fx ' and ' $p \ \& \ Fx$ ' are true only of existents, and whatever satisfies one satisfies the other. This uniqueness of truth-value making is contingent – not analytic – , hence it is a uniqueness pertaining to congruence (C), not to strict identity ($=$). When ' Fx ' is a contingent matrix, if ' $\iota x(Fx)$ ' does not denote, ' $F\iota x(Fx)$ ' is false, and so is ' $F\iota x(Fx \ \& \ p)$ '; otherwise, both denote and are contingently the same: that is, $\iota x(Fx) \ C \ \iota x(Fx \ \& \ p)$.

Since by hypothesis ' p ' expresses a truth and A exists, the following would seem to be true for the 1947 *Pseudo-Quine*:

$$(10) \ A \ C \ A \ \& \ p;$$

$$(11) \ (x)((Fx) \rightarrow (E!A \rightarrow FA));$$

$$(12) \ p \rightarrow (x)(Fx \leftrightarrow (Fx \ \& \ p));$$

$$(13) \ p \rightarrow (E!\iota x(Fx) \rightarrow (\iota x(Fx) \ C \ \iota x(p \ \& \ Fx)));$$

$$(14) \ p \rightarrow (y)(y \ C \ \iota x(p \ \& \ (x = y)));$$

$$(15) \ A \ C \ \iota x(p \ \& \ (x = A));$$

$$(16) \ \sim(A = \iota x(p \ \& \ (x = A)));$$

$$(17) \ N\sim(A = \iota x(p \ \& \ (x = A)));$$

$$(18) E! \ulcorner x(Fx) \rightarrow F \ulcorner x(Fx);$$

$$(19) N\sim(\exists x)(Fx) \rightarrow N\sim E! \ulcorner x(fx).$$

From Quine's premises *Pseudo-Quine* does not conclude that there is a vicious iteration of the Morning Star/Evening Star paradox. He concludes, rather, that the iteration shows that there are more distinct fine-grained objects than he may have initially thought there were. There are just more contingencies to be discovered.

II.7. Aristotelian essentialism

Quine applies the main idea of his argument quoted above in full to substantiate his charge that modal logic is committed to an incoherent view, which he calls "Aristotelian essentialism" (Quine 1961, 155; the italics express my different emphases):

Evidently this reversion to Aristotelian essentialism (cf. p. 22) is required if quantification into modal contexts is to be insisted on. *An object, of itself and by whatever name or none, must be seen as having some of its traits necessarily and others contingently, despite the fact that the latter traits follow just as analytically from some ways of specifying the object as the former traits do from other ways of specifying it.* In fact, we can see pretty directly that any quantified modal logic is bound to show such favoritism among the traits of *an object*; for surely it will be held, for each thing x , on the one hand that

$$(20) \text{Necessarily } (x = x)$$

and on the other hand that

$$(21) \sim \text{necessarily } [p \ \& \ (x = x)],$$

where ' p ' stands for an arbitrary contingent truth.

Obviously *Pseudo*-Quine is unhappy with this argument. He asks: “What are those objects, the values of the quantifiable 'x', Quine 1961 is talking about?” Blandishing Quine 1947's metatheorem for the semantics of modal logic, he is inclined to answer: “Those objects cannot be other than the fine-grained values of the quantifiable modal 'x'.” The trait being self-identical $[x = x]$ is possessed even by non-existing guises; the trait being self-identical it being (contingently) the case that p $[p \ \& \ (x = x)]$ is possessed by existing guises only, provided that p . Clearly, the guises $\iota x(x = x)$ and $\iota x(p \ \& \ (x = x))$ both fail to exist, because the uniqueness condition of individuality cannot be met. Suppose an individual a exists, that is $a \ C \ a$; then for any true proposition p , then there exists the different individual $\iota x(p \ \& \ x \ C \ a)$. Of course, $\iota x(p \ \& \ x \ C \ a) \ C \ a$.

Likewise, Quine's predicate matrices (20) and (21) yield two new non-existing guises. On the other hand, Quine's sentential matrices (20) and (21) yield truths for any value of 'x' – even if 'p' stands, alternatively, for an arbitrary falsehood.

Quine and *Pseudo*-Quine both, as noted above, construe predication as existential. This, as we have also remarked above, raises problems of interpretation. An object a is-essentially self identical, regardless of its existence; if it exists it is-existentially, contingently, self-identical. This is what $a \ C \ a$ expresses. Of course, $\sim N(a \ C \ a)$; likewise $\sim N(p \ \& \ (a \ C \ a))$.

Guises have-internally all their properties necessarily; existing guises have-existentially all their properties whatever. Thus, the fine-grained objects of modal logic shun Aristotelian essentialism. Let W be a coarse object, i.e., a system of our fine-grained guises. What is it for W to exist? what is it for W to have properties? W exists if and only all its constituents exist: W is a system of guises g and g_i such that $g \ C \ g_i$. So far neither *Pseudo*-Quine nor his possibilist counterpart have explicated what it is for W to have a trait or property F -ness. Undoubtedly, either such predication is senseless or it is analyzable in terms of F -ness a property constituting a guise composing W , for example:

(Pred.W) W is F : $W(F) = \text{Def. } (\exists g)(g \text{ is in } W \ \& \ g \ C \ \iota g(\dots \ \& \ Fh \ \& \ \dots))$.

(Pred.W.P-Q) $W(F) = \text{Def. } (\exists x)(x \text{ is in } W \ \& \ Fx)$.

We may certainly say that the singular terms, definite descriptions, which by Quine's 1947 metatheorem must denote guises, are derivatively ways of referring, in the case of existing guises, to their encompassing coarse C -systems. Then such systems have – in the sense defined – *all* their traits contingently, too.

Again, we may introduce a non-existential form of predication, to wit:

(EPred.W) $W \text{ is-e } F: W[F] = \text{Def. } \text{tg}(\dots \ \& \ Fh \ \& \ \dots) \text{ is in } W$. In this sense: Every coarse object has all its properties necessarily.

To conclude, if *Pseudo-Quine* were to extend his existentialist modal logic to quantification over coarse objects, which he conceives as systems of fine-grained guises, he would claim against the 1961 Quine that such massive objects have all their traits contingently. His Leibnizian possibilist counterpart claims, more comprehensively, that there are two forms of “connection of subject and predicate”,⁴ that in one form, $W(F)$, such massive objects have all their traits contingently, but in the other form, $W[F]$, such objects have all their properties necessarily. Therefore, neither the quantified modal logic proposed by *Pseudo-Quine* nor its possibilist extension countenances Aristotelian essentialism.

II.8. Pseudo-Quine's ontology and other sameness relations

Perhaps *Pseudo-Quine* can prevent the 1961 Quine's bomb from destroying his existentialist modal logic. Perhaps, as suggested above, his logic can be extended to a rich possibilist modal logic. Such an extension of quantification seems urgent because of the psychological modalities. Undoubtedly, we think of things that turn out not to exist, indeed some of them are even contradictory. There is, however, some doubt about the sufficiency of the contrast $\neq C$ to handle all the subtleties and limitations of thinking. Perhaps the psychological modalities require, not (yet) finer-grained individuals than quantified alethic modal logic, but only finer-grained

identities or *sameness* relations, or additional fine-grained objects of the same rank. Let's consider some types of experience which may require this rich program

A. *Necessary sameness and identity*. Recall that according to Quine's 1947 metatheorem distinct individual constants denote different fine-grained objects, and that *Pseudo-Quine* has taken definite descriptions as his primitive constants. Thus, $\sim(\ulcorner Fx \urcorner = \ulcorner Fx \ \& \ L \urcorner)$, where L is a logically true sentence (proposition); yet $\ulcorner Fx \urcorner$ is somehow necessarily the same as $\ulcorner Fx \ \& \ L \urcorner$. Patently, this necessary sameness does not depend on whether $\ulcorner Fx \urcorner$ exists or not.

Perhaps alethic modal logic might set aside their difference and conflate all those fine-grained individuals under the relation $N(\ulcorner Fx \urcorner = \ulcorner Gx \urcorner)$, if $Fx \leftrightarrow Gx$ is a logical (or analytical) equivalence. But this won't do for the psychological modalities. Consider the case of Diego Columbus, who has begun to study square roots. He has to learn that $\text{sqr}(361)$ is [the same as] 19. To be sure, 19 is necessarily [the same as] $\text{sqr}(361)$). Nevertheless, Diego does not have to learn that 19 is [the same as] 19, or even that $\text{sqr}(361) = \text{sqr}(361)$. He has to learn that the number that comes after 18 in the natural order of counting is the same as, converges with, the number delivered by *other* operations: root squaring. As Quine has pointed out all along, modal logics hypostatize the ways of referring to something as unique individuals. Hence the numbers Diego will find converging are for him distinct individual numbers. Since this convergence is necessary, it seems as if for a full understanding of the psychological modalities, we need, in the spirit of the 1947 *Pseudo-Quine*, distinguish between a necessary congruence, *conflation*, $*C$, and strict identity. Strict or genuine self-identity, $a = a$, continues to be necessary. This sameness is absolutely valid as a source of substitutions in all modal contexts; it is automatically learned whenever somebody learns something about an entity. However, $19 *C \text{sqr}(361)$ has to be learned separately, and yields no universally valid principle of substitution.

B. *Existential sameness and thought-of sameness*. Consider *The swamp*, a marvelous satiric novel that exposes the cruelties and tortures of the Pinochet tyranny in Chile. Its main character is an ugly voracious two-headed crocodile named Mata-todo. The effect of *The Swamp* hinges on the premise that:

(22) General Pinochet is [the same as] Matatodo.

Is this a different sameness, say, C^{**} , that *consociates* two fine-grained individuals, a contingent, writer-made sameness, that lacks existential import at least on one side? Is Sancho Panza C^{**} Don Quijote's squire, even though, $\sim(\text{Sancho Panza} = \text{Don Quijote's squire})$?

It may be said that literary characters have some sort of existence within literary works. Hence, perhaps, no special literary sameness is needed; perhaps the old sameness relations within the scope of literary modalities will do. Perhaps.

Nevertheless, artistic experience is rich in perplexities about sameness. Just one simple example:

(23) Marguerite Gautier has evolved: she has more depth in Verdi's opera *La Traviata* than SHE has in Dumas Jr.'s original novel *La dame aux camélias*.

The capitalized 'SHE' in (23) expresses a sameness across the two works of art, as follows:

(24) Marguerite Gautier of *La dame aux camélias* is the same as Violetta of *La Traviata*.

This sameness is a sort of *transconsociation*, say T^{**} , in the literary logical space *in between* the stories.

C. Perceptual sameness and individuals. Raymund and his visiting friend Erika are jogging in a forest in a foggy muggy morning. After a turn on the path Raymund points towards the bottom left of the view in front of them and asserts:

(25) That small gray flat triangle is [the same as] City Hall.

Erika looks attentively, and, remembering her tour of the city the day before, quickly responds:

(26) But City Hall is a huge perfectly cubical blue building.

Raymund acknowledges the puzzlement:

(27) I know; nevertheless in the winter mornings City Hall looks like *that* from here – I am not a painter, and don't know why.

The sameness posited by Raymund in (25) is very peculiar. I have called it *noncommunitarian* (in Jacobi-Pape 1990 and Gombocz 1989): not all the properties of the items flanking it transfer across it. As Erika's statement (26) shows City Hall is neither flat nor triangular nor gray nor small; likewise, the seen triangle is neither huge nor cubical nor blue. What sort of sameness is it, then?

The seen triangle is an item in Raymund's and Erika's visual fields. It is an indexical visual guise, which is a visual presentation that represents City Hall. Thus, the sameness posited in (25) looks like a new transcategorical congruence, CS^* , between a perceptual presentation and the physical item it represents.

Now, is there just one gray triangle that is seen by both Raymund and Erika? Or does each one have his/her own seen triangle in his/her own visual field? They concur in seeing a gray flat triangle; this, however, can be understood distributively. The intersubjectivity of seeing the same thing may be accounted for by positing parallel causal contexts which deliver similar perceptual contents. In any case, here we are just interested in raising questions about sameness. The outcome is that *if* seen contents are personal (or private), then there are perceptual fine-grained individuals to be reckoned with in the theory of perception. If such is the case, we also need a *transpersonal* sameness between subjective indexical fine-grained individuals to account for the sameness across their visual fields, thus:

(28) *That [Raymund's] small flat gray triangle C^{**} that [Erika's] small flat gray triangle.*

There are more questions about indexical items; we must forgo them.

III. Conclusion

The 1947 Quine brought forth the 1947 *C-* or *Pseudo-Quine*. This has advanced a defense against Quine's 1961 bomb. Nevertheless, the explosion of Quine's bomb has created a major problem of ontological reconstruction, especially in the development of quantified modal logics that can apply to the different spheres of human experience. Some such applications seem to require new species of fine-grained individuals. More importantly, they seem to require new sameness relations. Modalities, especially the psychological modalities, are logico-ontological prisms that break ordinary objects into a spectrum of fine-grained guises. Appropriately, they also break the old extensional identity into a manifold of sameness relations. Perhaps different types of experience can be explicated as pivoting on some peculiar and characteristic sameness relations determining appropriate massive objects from special domains of guises. The unity of experience can, perhaps, be accounted for in part by positing the total domain of guises where no extraneous sameness component is left in self-identity.

This is a very proliferative picture. In it all ways of referring to something have been hypostatized, and the ways of mingling those hypostases have multiplied. Further, as we discussed briefly: different forms of predication are suggested by the ways in which those sameness relations connect fine-grained objects. No entity without identity, certainly; also: no entity without entification, that is, without predication that makes it what it is. All of these issues should be part of the theory of individual guises.⁵

Is this the culmination of the *reductio* of modal logic Quine had envisaged? Indeed, how many sameness relations are too many epicycles? Which domain of fine-grained individuals, which sameness relation, will be the straw that breaks the back of guise theory?

Notes

An earlier shorter version of this essay was read and copies distributed at the Conference on Willard V.O. Quine's Contribution to Philosophy promoted by

Umberto Eco and his Centro Internazionale di Studi Semiotici e Cognitive at the University of San Marino, which took place in San Marino during May 21–26, 1990.

¹ Russell's Robust Sense of Reality is a weakened version of Parmenides's view that we can think and talk only about what exists. That version was held in the middle ages in an existential view of predication. William of Sherwood, for example, held it. See Jacobi 1980, pp. 318ff. For an intriguing reinterpretation and defense of the view see Kapitan 1990.

² For the central role of definite descriptions, which are the genuine rigid singular designators of a quantified modal logic, see Appendix.

³ The need to distinguish an existential from an essential form of predication was further strengthened in my mind by studying an excellent letter by William D. Hart announcing comments on the first version of this study, the one read at the Quine Conference. I have found the idea in Leibniz's *Discourse* Section 13; indeed, as I read his *Generales inquisitiones* 144 he is trying to stipulate canonical copulae for the distinction between existential and essential propositions. See Castañeda (forthcoming). I must confess that my reading of Leibniz's text was guided by the previous development of guise theory.

⁴ This is a translation of Leibniz's key phrase in his account of the difference between contingent and necessary propositions in his *Discours*, section 13:

Pour y satisfaire solidement, je dis que la connexion ou consécution [du sujet et du predicat] est de deux sortes, l'une est absolument nécessaire [...]; l'autre [...] est contingent en elle même.

⁵ The idea of guise theory preceded my study of Quine 1947, but guise theory was developed under the inspiration furnished by Quine 1947. That inspiration revolved around three elements: (i) the $C/=$ contrast, which yielded the multiple sameness/identity contrast network; (ii) the semantic metatheorem that quantified modal logic needs a domain of fine-grained individuals; (iii) the implicitly hinted at intimate fundamental connection between those fine-grained individuals and individual constants that denote them necessarily – yet what Quine called Aristote-

lian essentialism should be avoided. *Pseudo-Quine's* modal logic and its possibilist extension as mediating links between Quine 1947 and guise theory are retroactive developments on the occasion of Umberto Eco's retrospective San Marino Conference in honor of Quine. For exposition, discussion, and further development of guise theory see Castañeda 1982, Castañeda 1989, Tomberlin 1983, Tomberlin 1986, and Jacobi-Pape 1990.

Appendix

Quine 1947's contribution to the semantics of modal logic made essential use of individual constants or "names" that were really definite descriptions. As his project was polemical, pursuing a kind of *reductio*, he did not explore the nature of the individual constants on which his argument against modal logic depended. As we have seen *Pseudo-Quine*, engaged on a constructive program, could not fail to tackle the problem. He sees that quantified modal logic needs individual constants, furthermore, constants that capture individuals necessarily. *Pseudo-Quine* fully endorses Quine 1947's claim that the individuals of modal logic must be fine-grained. Thus, he has only his existentialist qualms in combining the ontological and the ontologico-linguistic thesis: the primitive essentially characterizing individual constants are definite descriptions. *Pseudo-Quine's* possibilist heir has no qualms at all in having his definite descriptions function as individual constants that necessarily denote their fine-grained denotata. For him nonelliptical definite descriptions are the required rigid designators of modal logic.

At the San Marino Quine Conference Paolo Leonardi and Ernesto Napoli presented a very interesting paper arguing against Quine's reduction of proper names to predicates. See also Leonardi (forthcoming). By *Pseudo-Quine's* lights, as well as mine, the Leonardi-Napoli's arguments are directed to two conflated questions, which must be separated:

- (A) Does human experience need singularly referring terms – which might be called individual constants?

- (B) If the answer to (A) is affirmative, are such constants *ordinary* proper names?

We (*Pseudo-Quine* and I) agree with them on the affirmative answer to (A); but partially side with Quine and disagree with them on the answer to (B). An attentive examination of how ordinary proper names really function in experience reveals that they are special common nouns. See Castañeda 1989, Ch. 2, and Jacobi-Pape 1990, Ch. 1. We also disagree with them and with Quine in taking nonelliptical definite descriptions as individual constants, and as not to be analyzed in the Russellian way. I also take single uses of demonstratives and other indicators as primitive individual constants.

Dagfin Føllesdal explained at the same Quine conference that, having been persuaded by Quine's animadversions against modal logic, since the early 1960's he saw that the only valid response to Quine's charge of Aristotelian essentialism was to assign to the objects of modal logic fixed names in all possible worlds. This amounts to taking proper names to name their nominata necessarily, which nominata are presumably coarse objects, rather than fine-grained guises. Here again we must distinguish two issues:

- (A*) Do Quine's arguments require that the values of the variables of quantification be necessary denotations of singular terms?
- (B*) If the answer to (A*) is affirmative, must such necessary terms be really proper names?

Pseudo-Quine and I agree with Føllesdal on the affirmative answer to (A). We disagree with him on his answer to (B): proper names are not the modally required singular terms. A proper name is characteristically empty of predicational content; hence it lacks the wherewithal to stick necessarily to what it names. A property of the form *being an N*, for a proper name *N*, is man-made, invented by the acts of naming objects *N*. (Not for nothing are proper names the paradigm predicates of nominalistic views.) *Pseudo-Quine* and I further agree against Føllesdal and Quine

concerning the Russellian analysis of definite descriptions. We take them as the individual constants that Quine's arguments require for modal logic to have constants that denote necessarily. They have contents that necessarily pick up fine-grained objects – the objects required by Quine's 1947 metatheorem. As explained in Sections I.7 and II.7, a Quinean modal logic – whether *Pseudo-Quine's* existentialist logic or its possibilist extension – is not committed to Aristotelian essentialism.

References

- Castañeda, Hector-Neri, *Sprache und Erfahrung* (Suhrkamp, Frankfurt am Main, 1982).
- Castañeda, Hector-Neri, *Thinking, Language, and Experience* (Minneapolis, University of Minnesota Press, 1989).
- Castañeda, Hector-Neri, "Leibniz's View of Contingent Truth in the Late 1680's" in Lenzen, Wolfgang, *et al.*, eds. *Festschrift* in honor of Heinrich Schepers (forthcoming).
- Cohen, R., *et al.*, eds. *Essays in Memory of Imre Lakatos* (Dordrecht, D. Reidel Publishing Company, 1976)
- Gombocz, Wolfgang, *et al.*, eds., *Traditionen und Perspektiven der analytischen Philosophie* (Vienna, Hölder-Pichler-Tempsky, 1989).
- Jacobi, Klaus, *Die Modalbegriffe in den logischen Schriften des Wilhelm von Shyreswood* (Leiden-Köln, E. J. Brill, 1980).
- Jacobi, Klaus, and Pape, Helmut, eds., *Das Denken und die Struktur der Welt, Thinking and the Structure of the World* (De Gruyter, Berlin, 1990).
- Kapitan, Tomis, "Preserving a Robust Sense of Reality," in Jacobi & Pape.
- Leibniz, Gottfried Wilhelm, *Discours de métaphysique* (1686), in Schepers 1989: 1687–1739.
- Leibniz, Gottfried Wilhelm, *Generales inquisitiones de analysi notionum et veritatum* (1686), in Schepers 1989: 1953–96.
- Quine, Willard Van Orman, "The Problem of Interpreting Modal Logic," *Journal of Symbolic Logic* 12 (1947): 43–8.

- Quine, Willard Van Orman, *From a Logical Point of View* (New York and Evanston, Harper Torchbooks, 2nd. revised edition, 1961).
- Quine, Willard Van Orman, "Whither Physical Objects?," in Cohen 1976: 497–504.
- Russell, Bertrand, "Meinong's Theory of Complexes and Assumptions," *Mind* 13 (1904): 204–19, 336–54, 509–24.
- Russell, Bertrand, "On Denoting," *Mind* 14 (1905): 479–93.
- Russell, Bertrand, Review of Meinong, *et al.*, *Untersuchungen zur Gegenstandstheorie und Psychologie*, *Mind* 14 (1905): 530–38.
- Schepers, Heinrich, ed., *Gottfried Wilhelm Leibniz: Vorausedition* (Münster, Leibniz-Forschungsstelle der Universität Münster, Faszikel 8, 1989): 1687–2056.
- Sharvy, Richard, "Three Types of Referential Opacity," *Philosophy of Science* 39 (1972): 153–71.
- Tomberlin, James E., ed., *Agent, Language, and the Structure of the World* (Indianapolis, Hackett Publishing Company, 1983).
- Tomberlin, James E., ed., *Hector-Neri Castañeda* (Dordrecht, D. Reidel Publishing Company, Profiles vol. 6, 1986).

CHAPTER 2

SUPERVENIENT PROPERTIES, EMERGENCE, AND THE HIERARCHY OF CONCRETE INDIVIDUALS

Introduction: Property supervenience, reduction, and emergence

Reductionist metaphysical claims are on the wane; but reductionism is not vanquished. Most philosophers nowadays recognize some degree of autonomy in different types of discourse, not only moral, legal, and literary discourses, but also in different species of scientific discourse. Some proclaim that there is no unity of science and no reduction of, say, biology to chemistry, or of chemistry to physics, let alone sociology to psychology or biology. Nevertheless, a unified ontological view that all reality is really physical is widely assumed, or explicitly adopted. This feels reductionist. To allay the tension between the two positions certain conceptual equipments have been used, all built upon a basic distinction between meaning and truth determiners. One such conceptual equipment is *supervenience*.

The main idea and the terminology goes immediately back to the British Moral Intuitionists in the first third of this century. Their leader G. E. Moore, for instance, claimed that, even though intrinsic goodness is an irreducible and simple quality, the intrinsic value (or goodness) of an object issues from (supervenes upon) the intrinsic nature of the object¹, thereby giving the pun a deep significance. Here we have the distinction between, on the one hand, the meanings of the ethical term 'good' and of the sentences of the form "X is good," and, on the other hand, the basic truth determiners of the propositions expressed by such sentences.

For the Intuitionists ethical supervenience was a logical implication of sorts, sometimes even called "entailment"; it was a synthetic *a priori* connection. This is

insightful, yet with the ensuing disrepute of the synthetic *a priori*, instituted by the Logical Positivists' semantic verificationism, the Intuitionists' notion of supervenience receded from the philosophical stage. Nevertheless, theses applying related concepts of reduction of truth but not of meaning have been propounded all along. Just two recent outstanding examples are Wilfrid Sellars and Donald Davidson². However, the vogue of the terminology and the promotion of such concepts to a field of study is a recent development, Jaegwon Kim being its chief theoretician³.

We must, of course, distinguish between the concepts of supervenience and the *theses* of supervenience, namely, the diverse claims to the effect that the properties (or phenomena) of a certain type supervene, in one or another sense, on the properties (or phenomena) of another type. It clearly is an open matter for investigation whether all reality is in some sense determined by or mappable on the physical, whether, for instance, the mental supervenes on the physical, and everything else on the mental, and by transitivity everything supervenes on the physical, or whether the mental and everything else is mounted on the physical in other ways than by supervenience. These questions can be woven in programs for empirical research that guide scientific progress.

This is a foundational study concerned with crucial ontological questions about the irreducible component in each case of supervenience. As we shall see reductionism is not entirely ruled out by supervenience. For one thing, there are reductionist and non-reductionist uses of the concepts of supervenience. For another, the schema of supervenience has a built-in reductionist strand: all subjects of predication are basic individuals upon which properties of different types supervene.

Evidently, the unity of the world and the unity of the experience and study of the world require that everything be, somehow and perhaps in different ways, connected to everything else. Whatever order reality has and can be apprehended must be at least partially manifested in the common world, and this is at bottom the physical world. Some primacy of physical truth is obvious. The nontrivial questions are how and to what extent physical truth determines all other types of truth.

Here we focus on the nature of supervenience determination. To inquire into the nature of truth determination, we must take a look at some phenomena for the

illumination of which supervenience has been propounded. To anticipate, that look and that inquiry, here cursorily reported, reveal that typically we confront, not merely the supervenience of properties, but the *emergence* of *Gestalt* properties or structures as possessed by individuals of higher types themselves *emerging* on the pedestal of individuals of less complex types. The problem of supervenience becomes, thus, the problems of elucidating, and accounting for, this *pedestal* relation between different categories of individuals. Yet one thing is transparent: unless one stipulates that complexity is irreality, that the complex does not exist, the claim of complexity is, rather, inimical to reductionism. Why should we make that stipulation?

In this preliminary study we just demarcate the problems. The demarcation opens up two complementary vistas. On the side of determination, it suggests that the truths of the pedestal categories by themselves may be only partial determinants of the emergent categories: perhaps everything is physical, except the structures of physical entities. On the side of non-reductionism, it urges an alternative to the flat ontology of supervenience, to wit: an Aristotelian-like hierarchical ontology of form *cum* matter.

1. Supervenience

In his fascinating paper Kim formulates several concepts of supervenience and explores their relationships. These concepts are all *holistic*: they pertain to classes, characteristically infinite, of properties as units of supervenience; they contrast with *atomistic* concepts that apply to single properties in isolation. Besides, their holism is structured. Supervenience applies to classes of properties closed under complementation (negation), conjunction, and disjunction: they are the Boolean closures of some base or generating class of properties. I will call them *Boolean* classes. Kim argues that the several concepts he discusses are equivalent to concepts satisfying one or the other of the following two schemata:

Weak supervenience:

(Kw) A Boolean class A of properties weakly supervenes upon a Boolean class B of properties, if and only if: NECESSARILY for any property F -ness in A : if an object x is F , then there is a property G -ness in B such that both x is G and every object y that is G is also F . (See p.163.)

Strong supervenience:

(Ks) A Boolean class of properties A strongly supervenes upon a Boolean class of properties B , if and only if: NECESSARILY for any property F -ness in A : if an object x is F , there is a property G -ness in B such that both x is G and necessarily every object y that is G is also F . (See p. 165.).

Remarks:

1. The two occurrences of the adverb 'necessarily' need not denote the same modality. Because the modalities are not specified the above biconditionals are schemata.

2. The distinction between the abstract individual F -ness and the predicative aspect *is F* is my own.

3. Let's call class B the *supervenience base*.

4. Because A and B are Boolean classes, the property G -ness may always be taken to be a maximal conjunction $(g_1) \& (g_2) \& \dots$ where each conjunct (g_i) is either one of the generating properties of B or its negation, and all the generating properties of B or their complements are included. Let us call such maximal conjunctions *Boolean conjunctions*. In his proofs Kim does use the assumption that G -ness is a Boolean conjunction.

5. The chief ontological difference between these two concepts of supervenience is that strong supervenience does, whereas weak supervenience does not, deliver logically valid or nomological generalizations linking the supervenient A -properties to necessary and sufficient B -conditions. This is important for the assessment of

several programs concerning the relationship between mental and physical statements (as Kim discusses in his paper, pp. 171ff).

From the above definitions Kim derives (pp. 169f) these results:

(Kw*) If A weakly supervenes on B , then for each property F -ness in A there is a property G^* -ness in B such that: anything is G^* if and only if it is F .

(Ks*) If A strongly supervenes on B , then for each property F -ness in A there is a property G^* -ness in B , such that: *necessarily* anything is G^* if and only if it is F .

In (Kw*) and (Ks*) the property G^* -ness involved is either a maximal Boolean conjunction (as characterized above) of base properties of B , or a disjunction, hereafter called *Boolean disjunction*, of such conjunctions.

2. The irreducible component of supervenient properties: Some questions

According to the supervenience theses, the supervenience of phenomena of type A on phenomena of type B involves the double claim that: (i) the truths in domain B determine the truths in domain A , so that, in the sense pertaining to the determination in question, there really are no A -truths; yet (ii) there is no reduction of the A -truths to the B -truths.

Palpably, supervenience claim (ii) is the contention that, in a sense pertaining to the type of reduction *not* accomplished, there is in each A -truth an irreducible residue. This raises some most important questions, among which lie:

- (Q*.1) What are the residues contained in A -truths [falsehoods] that can be thought of and are irreducible to B -truths [falsehoods]?
- (Q*.2) If truth and reality are exhausted by the B -truths, then what is the irreducible residue of an A -truth? Can that residue still be a truth component?

(Q*.3) Why do we need fully supervenient *A*-truths? Are they merely illusory, a fantasy built upon the *B*-truths on which they supervene?

These questions cry out for answers. Patently, a thesis of supervenience owes an account of the *irreducible* component of supervenient properties it is said to respect.

Some supervenientists are physicalists. They proclaim that because of the determination of every non-physical truth by physical truths, *everything is really physical*. Now, this crude physicalism looks reductionist. To it question (Q*.2) applies with the utmost urgency. Moreover, a philosopher who takes seriously that crude physicalism faces the temptation to adopt a further view – hereafter called *referential supervenientism* – that the denotations or referents of the *A*-expressions are really the denotations of the *B*-expressions with which they are paired by (Ks*) or (Kw*). This view looks even more reductionist yet. It provokes the following pressing question:

(Q*.4) What is the surplus, irreducible meaning of an *A*-expression whose reality referent is fully determined by the referents of the *B*-expressions?

Crude physicalism and referential supervenientism have to explain what they mean by non-reduction. Otherwise, their insistence on non-reduction is an empty verbal concession.

Supervenientism does not, however, have to be, or even appear to be, reductionist. Let us call *genuine supervenientism* any view that endorses either (Kw) or (Ks), shuns both referential supervenientism and the crude physicalistic claim. For genuine supervenientism, supervenient *A*-properties are truly irreducible. They have a standing of their own, and they contribute essentially to the *A*-propositions (i. e., *A*-truths or *A*-falsehoods). On this view typically the generalizations established by (Kw*) and (Ks*) are just hybrid *equivalences* between the supervenient *A*-propositions and the *B*-propositions in the supervenience base. Equivalence, however, is not identity.

3. The irreducible cores of supervenient properties: Equivalences vs. identity

The claim that supervenience does not yield a reduction schema must incorporate the thesis that logical equivalence – let alone empirical, nomological, conceptual equivalence – is not identity. This has always seemed to me an obvious truth, an obvious perplexing truth in that until recently it has been treated as unthinkable; its recognition is still not as widespread as it should. This obvious truth is, however, not the answer to questions (Q*.1)–(Q*.3); it is merely the gate to the path where the answers can be pursued.

Perhaps it is not amiss to establish that not even logical equivalence of the strongest sort can be identity. Consider any strictly monadic property *F*-ness, or *being F*, whatever, and take another property *Q*-ness, or *being Q*; consider now the disjunctive property *being F or (F and Q)*. The equivalence between *being F* and *being F or (F and Q)* is one of the strongest logical equivalences in the market. In general we can measure the comparative strength of a logical equivalence with respect to others by the kind of logical principles or rules on which it depends. The above equivalence depends merely on the rules of propositional logic and the principle that the logical connectives also apply to properties or to propositional functions. It is thus stronger than those logical equivalences that hinge on the preceding rules and the specific implicational rules of the predicate calculus. In any case, the properties *being F* and *being F or (F and Q)* are certifiably different. First, they have different logical form; second, they have different extensions: the extension of *being F* is the set of singletons that are *F*: { $\langle a_1 \rangle$, $\langle a_2 \rangle$, ... }, whereas *being F or (F and Q)* has two different, although corresponding, extensions: (i) the set of duplicative triples satisfying it: { $\langle a_1, a_1, b_1 \rangle$, $\langle a_1, a_1, b_2 \rangle$, $\langle a_2, a_2, b_1 \rangle$, ... }, and (ii) the collapsed set of ordered pairs satisfying it of the form { $\langle a_1, b_1 \rangle$, $\langle a_1, b_2 \rangle$, $\langle a_2, b_1 \rangle$, ... }. Patently, in no way can we regard *being F* as reducible to *being F or (F and Q)*: on the one hand, the reduction would be circular, and, on the other hand, the equivalence is logically ineliminable. It is worth observing that here the cleavage between equivalence and identity is not epistemological or psychological: it is logical and ontological.

Obviously, non-logical equivalences make no claim to be considered as identities. Nomological equivalences clearly presuppose the non-identity of the equivalent terms. Causal equivalences in particular are paradigms of non-identity and non-reduction. Nothing (aside perhaps God, who is no part of the world) causes itself. The causal connection [*a fortiori*, the causal equivalence] between a cause and its effect emphatically puts *both* into existence [or non-existence] as logically and ontologically distinct.

In brief, the nomological dependence of a supervenient property upon a base property presupposes an *ontological autonomy* on the part of the supervenient property. That is why the necessity of a supervenience equivalence of the form $(y)(G*y \leftrightarrow Fy)$ falls much too short of being reduction. Kim correctly insists on this.

Many supervenient properties also possess an *epistemic autonomy*. (Perhaps some degree of epistemic autonomy of some of the properties of a certain category is characteristic of the accessibility of that category to experience.) For instance, the domain of mental properties includes some that are epistemically autonomous by being directly accessible, namely, one's own states of consciousness.

To sum up, the non-reducible cores of supervenient properties are precisely the peculiar and pure properties of the supervenient domain. Some supervenient properties are mixed, and, of course, their components belonging to the supervenience base are trivially reducible. Some supervenient properties are to some extent directly accessible in experience. The crucial ontological fact is that supervenient dependence is not identity.

4. Individuals and supervenient reduction

So far we have discussed the non-reductionist character of genuine supervenientism. It has to do with the connections between domains of properties. Let us turn now to the possessors of properties.

Consider a *sophisticated* physicalism that includes genuine supervenientism. It does not endorse the thesis that everything is *really* physical, but only the dependence-*cum*-autonomy thesis that every single property, whether we encounter

it in experience or not, is either physical or ultimately genuinely supervenient on physical properties⁴. This thesis is accepted by many philosophers.

Now, the holistic character of Kimian supervenience introduces a limitation. Kim has deliberately (p. 258) restricted his definitions of supervenience to non-relations. If this restriction excludes all relational properties, like being to the left of Archie Jones, or being twenty millions richer than John Rockefeller, then Kimian genuine supervenientist physicalism may not get off the ground. There may not be enough relevant physical properties to have any supervenience at all. The psychological properties are essentially relational: just consider the simple cases like believing *of John* that he is miserable, assuming that *the earth* is round, planning to attend *Smith's funeral*, etc. On the other hand, if all relational properties are included, then there is a question as to how many individuals can have the basic properties. By Leibniz's law every individual alone has a maximal Boolean conjunction of properties. Given supervenience, the supervenient properties cannot help to differentiate one individual from another. All the differences between individuals must lie at the basic level of physical properties. If this supervenience is Kimian, a maximal Boolean conjunction of all physical properties or its complements is the relevant basic property for supervenience. The supervenience conditions will then be met trivially by each individual. This raises a problem about change. The mixed *A-B* generalizations of (Kw*) and (Ks*) would be essentially vacuous – unless we adopt a substratist view of individuation and give up Leibniz's law.

The preceding comment is not an objection. The Kimian formulas can be revised, first, to include relational properties, and, second, by afterwards introducing a basic proper sub-class *B** of the class *B* on which the *A*-properties supervene. The preliminary point is merely to introduce the problem of individuation: *a usable concept of supervenience has to be mounted on a theory of individuation.*

It may be worth pointing out that even a substratist view of individuation has a hard time giving up Leibniz's law. Just to mention some trivial hurdles. Make the *assumption* that the universe is with two distinct but indiscernible objects, i.e., for any property *F*-ness: one is *F* if and only if the other is *F*. Then there is an ordered pair such that one has the property of being its first member but the other does not;

of course, there is another ordered pair such that the latter has the property of being its first member, whereas the former does not. In short, the two objects must differ in at least *two* properties. This contradicts the assumption of indiscernibility. Another hurdle is this: there has to be a distance d between the two objects, and one is at no distance from itself, while the other is d from it. Here again we have two properties in which they must differ⁵. Let's not dwell upon the fact that, although both share the property of being self-identical, there are two other relational properties in which they must also differ: one has the property of being identical with it, but not the other.

A look at definitions (Ks) and (Kw) above, or, more clearly yet, a look at theorems (Ks*) and (Kw*), shows a most important ontological commitment of Kimian supervenience. The object (individual) that possesses a Boolean conjunction of *B*-properties is exactly the selfsame object that possess the corresponding supervenient properties. Thus Kimian supervenience, if it obtains all across, delivers a uniform picture of reality: one single domain of individuals, with basic and supervenient properties, compose the whole of reality. But what are those individuals?

Consider again sophisticated supervenientist physicalism. What are the physical individuals that possess *all* the properties, basic and supervenient?

Physics studies a pretty large hierarchy of entities. And it is widely held that all physical entities in that hierarchy are reducible to the boot rank: domain of ultimate particles. Armed with this assumption, the fundamental (reductionist) physicalist adopts a rigorous *mereological* position: the real individuals just *are* the ultimate particles – or whatever – that physical theory will posit – when finished, of course!

In brief, the theses of supervenience are at bottom reductionist concerning the *possessors* of properties. Is this reductionist strand acceptable? What is its underlying grain of truth?

5. A look at the world

Perhaps in the case of axiological and deontic properties, the ones that concerned the Moral Intuitionists, it is the same domain of objects that have both the base *B*-properties and the supervenient *A*-properties. The same entity that is, say, an experience of pleasure is precisely what has some amount of intrinsic goodness; the very same person who tortures another for the sake of enjoying the sight of pain behavior is an immoral person, and it is the same act of torturing that is morally wrong.

Nevertheless, as a general conception of truth dependence, the extreme definitional reduction of the possessors of supervenient properties to the possessors of the base properties is out of order. Palpably, the psychological properties we have are not the properties of the ultimate particles to which all physical individuals are reduced. Perhaps quarks – to stay within current physics, without waiting for the end of physics – possess psychological properties, but the psychological states and episodes of the quarks composing your body are certainly not your own mental states or acts. Perhaps a mentalism of such quarks will be called forth by later physics, and perhaps the mental states of quarks are somehow at the basis of the mental states of the persons whose bodies are ultimately composed of such quarks. But this is much more problematic than the connections between mental powers and states of consciousness the neurophysiologists are searching for in the brain. Whatever dependence of the mental on the physiology of brains those scientists may establish or propose seems not to necessitate the attribution of mental states as such to neurons, let alone the quarks in those neurons.

Undoubtedly, the colors we see depend on complex situations involving our physiological condition, on the waves causing us to see, and the situation surrounding the objects emitting the light impinging on our eyes. But the light is not the subject of predication that has the colors, much less each of the photons of the light emitted by the colored surface.

A muscle is flexible, healthy, infected with this or that kind of virus, young or old, etc., but not the cells, much less the molecules, composing it, not to mention the ultimate particles of current physics.

Therefore, if we are to find a useful general conception of ontic dependence that can illuminate the nomological unity of the world, we must replace Kimian supervenience, because of its reduction of the possessors of supervenient properties, with a similar but more liberal concept that is even less reductionist. Let's call *emergent properties* those properties that accrue to a whole because it has certain parts and because of the interaction of these parts with its environment. Let's call such wholes *emergent individuals* or *objects*.

6. The hierarchy of individuals: The irreducibility of structure

The world has, it seems, unreduced possessors of emergent properties. Their ontological autonomy grounds the ontological (and epistemic) autonomy of the supervenient and emergent properties. Objects are embodied structures, and these structures are the ultimate irreducible components. These structures belong to levels of organization that presuppose embodiment in materials with their own structures. To firm this point let us consider other cases.

A. *Michelangelo's David*. This famous sculpture is made of marble. Yet the statue is NOT identical with its marble. They differ in the important property that the statue has the *power* to remain in existence and precisely where it is located while losing all its marble. This power can be exercised in many different ways. Here is one. The atmosphere in Florence becomes so polluted that the marble molecules of the sculpture have slowly been replaced with molecules of some other kind. After some years, the marble is, so far unbeknownst to all observers, all gone and some new material composes the statue. Yet the David looks just as marvelous as ever, with the same glorious combinations of lines, surfaces, volumes, colors, and texture.

To be sure, the statue is dependent on *some* material or other, but it does not require the material actually composing it at any given time. The statue is an individual that *emerges* upon the individual material composing it. Some of its sculptural properties emerge upon the properties of the material; but not all of them. The statue is a *materialized structure*, but different from its matter.

B. *Theseus ships: An enriched version of the hackneyed story.* For quite some time now every night that Theseus has moored his ship to a Cretan pier, one-eyed Polyphemus has been replacing one part (plank, bolt, or whatever) of the ship with a part that looks the same, is made of the different but sometimes similar materials, and functions sometimes equally well, but not always. As soon as Polyphemus leaves the cave where he deposits the fruits of his robbery, Leander of Sparta enters the cave with his new-fangled gadgetry and replaces the crystals composing the stolen part with crystals of similar type; he lays the original crystals next to a blueprint for reconstructing the original parts of the ship the next morning. But before the next dawn Alcibiades of Corynth enters Leander's cave with more sophisticated equipment, and replaces the molecules in Leander's crystals. In a cave further down the mountain road he stores his stolen property next to his plans for building up crystals and ship's parts some days later. However, before Alcibiades starts building, Euripides of Acadia enters Alcibiades's cave and replaces the atoms in the displayed molecules; he takes the original atoms of Theseus's ship to his own cave and lays them out in adequate patterns to build from them molecules, crystals, and ship's parts. But before he starts building, Sophocles of Thebes, who always waits for him to leave his cave, walks in and replaces the electrons and nuclei of Euripides's atoms. ... The chain of replacing, stealing, storing, and reconstructing the tampered wholes goes on down the row of Cretan caves till the smallest physical particles, the recently well established s3q's (in old jargon: sub-sub-sub-quarks), which Zeno of Syracuse steals, stores, and uses in building up the appropriate wholes. The theory positing s4q's is still being tested. In any event, one good day all of the above mentioned fiction-scientists put at once their replicas of Theseus's ship.

The dispute as to which one is the real Theseus's ship will most likely go on beyond the Day of Judgment. One answer, Locke's: the (ordinary) object Theseus's ship is the ship that Theseus or his successors operate, work on and in, repair, and rent: Theseus's ship is an *institutional* object: it has its own institutional unity and identity that undergirds its replaceable parts, and the replaceable parts thereof, and so on. The notion of (world) object is, as Protagoras, Locke, and Heidegger have emphasized, *anthropomorphic*: (world) physical objects are tools or potential tools

in human activities and have their identity through the roles they play in the human or thinkers' activities in which they are involved: their physicality is structurally necessary, but no physical individual is itself a necessary component. This Protagorean-Lockean-Heideggerian insight is valid even for objects illustrating so-called natural kinds: these are neither themselves metaphysically necessary nor have a non-anthropomorphic unity. This insight is what contemporary biological theory and its ensuing technology aim at making obvious.

The ships built by Polyphemus, Leander, Alcibiades, Euripides, and ... and Zeno are of course further institutions and tools. Conceivably the *institutional* unity of each includes *from above*, wholistically, that they be constructed in the way indicated and even that, if certain obtaining conditions are worth stressing, they be made up of the particular physical components they are in fact composed of. For instance, those replacements of parts could have been part of a large scientific research project, and the ships they built could have been further institutionalized as historical pieces in some museum for the history of science. Then the individuality of those physical objects *qua* physical individuals may be sanctioned.

C. *Persons and biological wholes.* Persons are possessors of psychological states and episodes of consciousness. Patently, they are not strictly identical with their bodies, just as much as their bodies are not strictly identical with their organs and limbs, just as much as their organs are not strictly identical with their composing tissues, just as much as tissues are not strictly identical with their component cells, just as much as these are not strictly identical with their component chemicals, just as much as these are not strictly identical with their component atoms, and these are not strictly identical with ... At every stage there are structures that maintain the unity of the entity from above in spite of the potential or actual changes of components. The nature of the entity consists of those powers and counterfactual properties that endure while components and materials are replaced or even become lost.

D. *Institutions and societies.* Sociological individualism (the special reductionist view of social groups to their members) has also the problem of individuation. Undoubtedly, the realization of an action by a social group or an institution is a

network of actions by certain members of the group or institution. These actions must of course be activity of some last particles of physics. Clearly, the attribution of mental states to institutions are metaphorical: they are the mental states of some members of the institution. On the other hand, the *normative* relations (that is, obligations and rights) of an institution are irreducible to normative relations of any of its members. An institution *I*'s obligation to do *A* implies, without reduction, obligations on certain officers of *I*, but the officers are replaceable and even the offices themselves may be modified. For instance, a company *C* has the obligation to pay a fine of \$6,000,000 to the government of the United States for fraud in some defense contracts; depending on *C*'s organization that obligation yields, say, an obligation on *C*'s treasurer to sign a check, an obligation on *C*'s financial vice-president to approve the check, and an obligation on *C*'s president to countersign it. If *C* undergoes reorganization these obligations may be transferred to other officers. Of course *C* has those obligations even if during the reorganization the financial vicepresident dies and the new office to be in charge of approving checks over \$5,000 is vacant.

7. Mereological essentialism and world-object contingentism.

The part/whole contrast is a serious source of puzzlement concerning the identity of objects. It engenders the tension between two opposed tendencies. On the one hand, there is a tendency to *mereological essentialism*: to view objects as essentially composed of their parts. That tendency does have *some* ground. Objects have certain overall or *Gestalt* properties because of the properties of their parts. Furthermore, for many different purposes we need to deal with the parts of objects, e.g., to replicate them, to create substitutes, to repair broken or dysfunctional objects. On the other hand, there is the tendency to *wholism* or *molarism*: to view objects as molar structures, which ground irreducible emergent properties, and within which properties and parts compose objects. Here the word 'wholistic', arising from the common word 'whole', contrasting with the Greek-originated 'holistic', signals that we are concerned, not with large and lofty metaphysical wholes, but with most vernacular wholes, which are far from constituting any exhaustive Whole.

The mereological/wholistic tension has become very intense because of the prestige and the success of physics. All branches of science to a large extent study the composition of different types of objects. But one part of physics is devoted to the postulation of simpler and simpler entities of which everything else in the world seems to be composed. The spectacular illumination wrought out by the positing of so many layers of micro-atomic particles has fed a natural inclination to understand wholes in terms of their parts, and this inclination has fostered the reductionist view that wholes are nothing but their parts and the relations among these, that wholes do not really exist. Here is much in need of clarification; yet I am prepared to accept:

(PME) *Physical Mereological Essentialism*: A physical object *qua* physical individual is essentially composed of its parts, that is: the replacement of a single physical component yields a different physical entity.

This is, however, compatible with treating the ordinary objects of the world as irreducible wholes, with substitutable parts, all of which, although at bottom physical, are not treated *qua* physical wholes. In fact, the question properly arises as to whether physics itself always, or ever, treats physical objects *qua* physical individuals. The extreme individualism of Mereological Essentialism is foreign to the theoretical goal of physics, namely, to deal with very general structures. In physics Mereological Essentialism has, at best, a place in experimentation. The other sciences, by dealing with less pervasive phenomena than the purely physical, take their stand at different levels of composition; these levels are constituted by the establishment of certain categorial forms under which substitutability of parts is, if not required, at least not denied. For instance, the new technologies being developed in biology aim at creating not merely new types of cells, but also exemplars of the old types not only with different biological components, but even with different chemicals. Thus, we must also adopt:

(W-OMC) *World-Object Mereological Contingentism*: The parts of an object *O* not considered merely *qua* physical individual are all inessential to *O* and

dispensable in *O*, i.e.: *Each* part *P* of *O* can in principle be replaced with a part *P** that discharges the same, or even just a similar, function *P* has in *O*. Furthermore, some world-objects may even lose parts and yet remain self-identical.

8. Conclusion

The chief moral is that the notions of *physical individual* and (*ordinary world*) *object* are organized in diametrically opposite directions. Physical individuals are built upwards from their components: the ultimate particles of physics are the ultimate constituents of the universe – regardless of there being human beings in it or not. The objects of the world, the objects we find or construct in *our* world, structure downwards from the unitary conception as wholes. As hinted at above, even the expression 'physical object', not really a term of physics, does not mean 'physical individual'. It is semiotically significant that physicists talk, rather, of *physical particles*.

Kimian supervenience is a narrow and useful concept. Extended to relational properties it seems adequate to explain the ontic dependence of normative and axiological properties on other types of properties that have the same subjects of predication. Yet more widespread is the ontic dependence consisting of the emergence of individuals with their proprietary domains of emergent properties. The principle of emergence yields a hierarchy of individuals composed of both an appropriate material and a characteristic structure. Even theologically-oriented philosophers may find sophisticated emergentist physicalism agreeable. Certainly Christian philosophers can rejoice at its echoing Aristotle's view of individuals as composites of a form having a hierarchical ontological level and of a matter suitable for the form in question.

Notes

An early version of this paper was presented at the meeting of the Deutsche Gesellschaft für Semiotik in Essen, October 1987.

¹ See G. E. Moore, *Principia Ethica* (Cambridge: Cambridge University Press, 1903), and “The Conception of Intrinsic Value” in *Philosophical Studies* (London: Routledge and Kegan Paul, 1922); William David Ross, *The Right and the Good* (Oxford: Clarendon Press, 1930). While Moore ends up making the value of a whole depend on the intrinsic nature of the whole, Ross maintains that it depends on all of the whole and called ethical properties 'toti-resultant'.

² Wilfrid Sellars developed early in the 1950's a combination of ontological physical reductionism with semantic autonomy. This became part of his contrast between the Manifest Image of man-in-the-world and the Scientific Image. See his “Philosophy and the Scientific Image of Man,” in Robert Colodny, ed., *Frontiers of Science and Philosophy* (Pittsburgh: University of Pittsburgh Press, 1962). Donald Davidson proposed a similar type of view in “Mental Events,” in L. Foster and J. W. Swanson, eds., *Experience and Theory* (Amherst: University of Massachusetts Press, 1979).

³ Jaegwon Kim, “Concepts of Supervenience,” *Philosophy and Phenomenological Research* 45 (1984): 153–176. This is the culmination of preceding essays.

⁴ Physicalism as here broadly characterized may be the physicalistic view advanced by Wilfrid Sellars (e.g., in “The Language of Theories,” in Herbert Feigl and Grover Maxwell, eds., *Current Issues in the Philosophy of Science* (New York: Holt, Rinehart and Winston, 1962). This is a matter of exegesis we cannot carry out here.

⁵ A celebrated attack on the identity of indiscernibles built on symmetric universes is Max Black, “The Identity of Indiscernibles,” *Mind* 61 (1952). The (alleged) counterexample of the universe consisting of two identical spheres one mile apart is Black's. A comprehensive discussion of the main elements of the debate is Gregory Landini and Thomas Foster, “The Persistence of Counterexample: Re-examining the Debate Over Leibniz's Law,” *Noûs*, forthcoming 1991.

CHAPTER 3

INDEXICAL REFERENCE AND CAUSAL DIAGRAMS IN INTENTIONAL ACTION

To deliberate about what to do is to engage in an investigation about one's own body and its future. To face a conflict of duties is on the surface to confront a conflict of reasons for doing this or that; but down deeply inside oneself it is for oneself, as agent, to live a conflict of bodily inclinations to do this or that. Reasons for action, taken *de re* are not causes, but inclinations to act. To solve the conflict is ostensibly to sort out the conflictive reasons in order to find out what one ought, everything considered, to do; but down deeply inside oneself it is to establish an order of causal dominance among the conflictive inclinations to act. The direction of that causal order is thought of as an intention to do. To deliberate is to gain reasoned beliefs about one's body. The totality of the actions one believes one can do compose one's picture of one's body.

ROD CHASTEIN

Consciousness is too complex and too precious a commodity to be squandered away. In an admirable universe it is not squandered. That is why we need consciousness and hard thinking when we are learning practices and acquiring habits. Thereafter, our behavior is habitual, with great economy of consciousness and great efficiency. Yet habitual behavior can be intentional.

OSCAR THEND

To come to know what to do is to have a thought which itself consists of an awareness of its bringing about an action, or a rearrangement of one's causal powers ... The causal dimension of practical thinking is the coalescence of contemplation and the causation of that contemplation, and the contemplation of that causation.

Thinking and Doing

Introduction: Four problems about intentional action

Here I am interested in investigating the general structure of intentional action that underlies and unifies the solutions to four exciting problems, namely:

- A. The intentionality of intentional action.
- B. The Aristotelian practical syllogism, characterized by having an action as “conclusion,” and the fact that the action is caused by the agent's thinking of the minor premise.¹
- C. The practical causality of practical thinking.
- D. The proximate causal antecedent of intentional action.

Topic D is a special case of what Myles Brand has called “the question in action theory,” which he has nicely formulated as follows:

What properties must a mental event have in order for it to be the proximate cause of action, or more generally, the proximate antecedent of action? A shorter, more general version of this question is: What initiates action?²

Brand's question is deliberately open with respect to the mind-body relationship. This is as it should be. However, we must remember that concerning action, a mind that acts in the world must be embodied; hence an embodied mind is tantamount to a minded body. But what is an *action*? Brand's question involves the semantic stipulation that a '(human) action' is a sequence of changes caused by a mental event

of some agent. This may be too narrow a sense of 'action'. For one thing, some changes caused by an agent may very well have their causality solely within the non-mental properties of events in the agent's body. For another, we should not rule out in advance that what we sometimes refer to as an agent's nonintentional and involuntary action is either not really an action or is caused by some mental events of the agent. Nevertheless, be this as it may, although not all human actions are intentional, intentional action is the obvious case of physical changes caused by the agent's mental states. Hence, Brand's question applied to intentional action is fine. Yet as we understand it we must *not* commit ourselves in advance to there being just one kind of mental event, whether this is reducible to, or is parasitic on, a physical event, as the cause of intentional action. In fact, to anticipate, we claim that mental events of different kinds can be the immediate causal antecedents of intentional action.

Problem D is a special case of problem C. Practical thinking includes not only the thinking of intentions, but also the thinking of obligations and duties, and the thinking of imperatives. As said above, here we are interested only in the general structure that underlies those problems. That structure reveals the unity of their solutions.

1. Voluntary and intentional action

Our major problem is the causation of intentional action. The central cases of intentional action are actions performed at will. This raises the question about the connection between voluntary and intentional action. To demarcate our problem let's start by elucidating this connection.

The adverb 'intentionally' places emphasis on what the agent thinks, i.e., in how she thinks of the action she is to perform and of how she conceives of the persons and objects involved in his/her action. The adverb 'voluntarily' emphasizes the mechanism of the agent's doing of her action. Let's reflect on the following situation:

- (1) Diana's vision became blurred and she wanted to see an oculist;
- (2) Diana intended to go to see her oculist;

- (3) Diana believed that her oculist was Dr. Zahn;
- (4) Dr. Zahn is not an oculist, but a dentist;
- (5) Diana chose correctly the way to Dr. Zahn's office;
- (6) Diana went to Dr. Zahn's office.

In this case it will be said that, since nobody forced or coerced Diana to go to Dr. Zahn's office in any way:

- (7) Diana went to Dr. Zahn's office voluntarily.

Also:

- (8) Diana went to Dr. Zahn's office intentionally.

From (4) and (5) it follows that

- (9) Dr. Zahn's office is the office of the dentist (of the story).

Therefore, from (7) and (9):

- (10) Diana went to the dentist's office *voluntarily*.

But because of (2)–(4),

- (11) Diana did *not* go to the dentist's office intentionally; she went there by mistake.

On the converse side, some people would contrast 'intentionally' and 'voluntarily' as in:

(12) Cynthia pointed a gun at Marcus and told him to give her the money bag in the cash register. Marcus resisted, but Cynthia cocked her gun and ordered Marcus again to give her the money bag. Marcus was frightened; slowly he took the money bag out of the drawer, and even more slowly he handed it over to Cynthia. Thus, Marcus gave Cynthia the money bag intentionally (with full awareness of what he was doing, having decided that his life was worth preserving), but he did not do it voluntarily.

There is a sanction for the contrast between 'intentionally' and 'voluntarily' recorded in (12). We have there the *non-coercion* or “*freedom*” sense of 'voluntary action'. On the other hand, some people would say:

(13) Since Marcus *chose* to save his life and hand the money bag over to Cynthia, he acted voluntarily: he did it at will.

In (13) we have the *willingness* sense of the words 'voluntary' and 'voluntarily'. The terminology is not important, the differences in reality are. The two senses of 'voluntary action' originate in different areas of discourse about human action. The non-coercion sense is relevant in normative contexts, when we are engaged in assigning responsibilities and handing down punishments, or rewards. In the example (12), Marcus will not be regarded as responsible by being robbed by Cynthia. On the other hand, the willingness sense of 'voluntary action' appears in discussions about the connection between will and action. Given our present interest on the latter topic, we will prefer the second sense of 'voluntary action'.

There is some *objectivity* in the attribution of voluntary action to an agent, whereas the attribution of intentional action is *purely subjective*: it depends on how the agent conceives his/her actions. The objectivity of voluntary action is, however, not perfect. The agent's willingness can only accrue to an action that he/she conceives in a certain way. I propose the following connecting principle:

(V*.A) “Agent *a* does (did, will do) *A* to entity *b* voluntarily” is equivalent to
 “There is a property *F*-ness such that: *b* = the only entity that is *F*, and *a*
 does (did, will do) *A* to the only entity that is *F* intentionally.”

With (V*.A) at our disposal we concentrate hereafter on intentional action understanding it in the willingness sense.

2. The intentionality of intentional action

A significant development in action theory is a growing tendency to abandon atomistic conceptions of action, in particular, of intentional action. Most of us agree – I believe – that an intentional action is characteristically or typically *not* an action that is merely done by an agent who has the intention to do it. For one thing, having an intention is a dispositional state and not an event. A person does not lose her intentions when, for instance, she falls asleep. Episodes of thinking are *mental events*. These include comings to intend to do some action *A* as well as thinking episodes in which the agent re-affirms her previously adopted intention to do *A*. Let's call events of these two sorts *rehearsals of intending* to do *A*, and, for short, *volitions* to do *A*. Since mental states and mental events are characterized by their contents, we must enquire into the contents of volitions. Then we can tackle the first step in dealing with problem D above: Is *X*'s volition (or rehearsal of his intending) to do *A* the event that immediately causes *X*'s intentionally doing *A*?

Just rehearsing his intentions is not sufficient for an agent to fulfill them intentionally. A person attentively and volitively thinking of what he intends to do may bring it about accidentally and unexpectedly. Thus, Jones, set on killing Smith, may go to Smith's house thinking all the way: “I'll kill you, bastard!” In the way Jones throws a rock to a tree to practice his aim, and lo! and behold!, unbeknownst to Jones, Smith is there and the rock hits him on his forehead wounding him mortally. Patently, Jones does not kill Smith intentionally. For an agent to do an action *A* intentionally, his intention to do *A* must, as the law textbooks say, “actuate” his doing *A*. But what is for an agent's intention to do *A* to actuate his doing *A*?

An intentional action is not merely an action one brings about by the causal power of one's rehearsal of one's intending to do it. Rather, *an intentional action is an action that belongs to an intended course of actions*. Thus, intentional action has a holistic or *Gestalt* nature.

Particularly important in this regard is a datum discussed by Gilbert Ryle³. He considers the case of a very skilful clown who performs a very complex sequence of movements for, say, 10 minutes. Ryle points out, *correctly*, that there is no reason to suppose that the clown has a volition for each movement. Indeed, it may be very hard to break down the performance into unitary movements in order to assign a volition to each movement. Perhaps the clown performs the whole sequence at will, the sequence of movements issuing just from one inceptive volition. On the other hand, perhaps the clown intersperses some volitions alongside his performance. Half way through the performance he may feel insecure, but he reassures himself by thinking "I must go on to the end." Then later he discovers himself weakening and gains back his strength to continue by means of another volition: "I'll finish! Regardless of the pain." The crucial point is, however, that some particular movements will be performed intentionally – and at will – , not because they are covered by their very own proprietary volitions, but because they are part of a train of movements that is *as a whole* performed intentionally.

Ryle's magnificent clown also shows another central fact: Actions themselves are not just bodily movements: movements are temporally and spatially divisible to infinity, or to the physical limits of such divisibility; human actions are not. Actions have a basic *Gestalt* character: they occupy intervals of space and time. This property enters crucially in intentional action. I intend to move from *A* to *B*, without intending to move to or through any given point *C* between *A* and *B*. Indeed, it does not even enter my mind that the distance between *A* and *B* is – or is not, for that matter – a continuum of points to be matched by the continuum that is the time in which I will traverse it. My moving from *A* to *B* may be as far as intention is concerned an indivisible act.

3. Action, doing, characteristic state

Before we investigate further the structure of intentional action we must elucidate what we mean by the fundamental term 'action'. As I have shown in detail elsewhere⁴, a huge amount of discussions about action, about the number of actions one performs at a given time with one bodily movement, about the identity of actions, about the time and place of action, and other topics are typical philosophical disputes in which the disputants do not join issue, but simply make some true claims and attack the true claims of their opponents after a tergiversation of these claims by using the key terms in different senses. Let's explain how we use the key terms in this study. Consider the following case:

The Tragedy of Francesca. Francesca di Verona has at 7 a.m. decided to kill her lover, Romeo di Calabria, who is no longer interested in her. She has decided to kill him; but she has no idea about the method of killing. Nevertheless, she feels relaxed after she reaches her decision and falls asleep. She wakes up at 11 a.m. She feels good. After her slow toilet she eats a hearty breakfast. At 1 p.m., while taking a bath she starts thinking about the way to kill Romeo. After some reflections, having discarded poisoning, strangulation, suffocation, pushing Romeo down a ravine, cutting his throat with her new kitchen knife that he bought for her, and other techniques, she sets upon shooting by gun. At 2 p.m. she buys a gun, and at 2:30 she enters Romeo's apartment. He is in the bedroom, she in the livingroom. At 2:31 p.m. the gun appears in her right-hand. At 2:32 +t the gun fires, and at 2:32 +t +t' p.m. the bullet hits Romeo. At 12 midnight Romeo dies in the kitchen.

Here we have the interplay of the following types or concepts of action:

I. Actions₁, or action-properties:

- a) *General*: shooting; suffocating; strangling, suffocating; poisoning; firing a gun; killing a man with a gun;

- b) *Object-particularized*: strangling Romeo; poisoning Romeo; cutting Romeo's throat; cutting Romeo's throat with a knife bought by Romeo; killing Romeo by firing a gun at him.
- c) *Agent-particularized*: Francesca shooting; Francesca poisoning someone; Francesca poisoning Romeo; Francesca killing Romeo by shooting him in the kitchen;

II. *Actions₂ or realized actions*:

- a) Buying a gun, flexing a finger, firing the gun, shooting Romeo, wounding Romeo, killing Romeo, etc.
- b) Francesca's buying a gun; Francesca's killing Romeo with a gun;

III. *Characteristic states of actions*:

- For the action₂ of (Francesca's) shooting Romeo: Romeo being one who was shot;
- for the action₂ of (Francesca's) wounding Romeo: Romeo being one who had a wound;
- for the action₁ of (Francesca) killing Romeo: Romeo being dead;
- for the action₁ of (Francesca) flexing her index finger: Francesca index finger being one that was flexed; etc.

IV. *The doing of all those actions*:

- Francesca's right index moving at 2:32 p.m.

V. *Actional existing states of affairs (true propositions), or facts*:

- Francesca flexing her right index; that Francesca shot Romeo; that Francesca fired the gun; that she killed Romeo;

Va. *Special temporalized facts*:

- Francesca flexing her right index at 2:32 p.m.;
- Francesca picking up her gun at 2:31;
- that Francesca fired her gun at 2:32+t p.m.; etc.

It is my contention that to understand intentional action we need yet another concept of action, which I call *practition*.

We must remember that there is a continuous flow of energy, and that the energy does not follow one path. The energy distributes in a region of varying contour and size. There are other effects from the explosion in the gun chamber. In describing the tragedy we have selected just one path of the flow of energy. This selection hints at a conventional, social dimension of our notion of action. (See Note 4.)

4. Goldman's analysis of intentional action: Praise and critique

Alvin Goldman, in a book⁵ I like very much, and with an emphasis I cannot stop admiring and commending, has provided a full explication of a concept of intentional action that places the *Gestalt* character of intentional action in high relief. Goldman builds his notion of intentional action on the notion of an action-plan, which is the most brilliant idea in his book. One cannot overemphasize the significance of treating wants, beliefs, and intentions as concerned with structures of possible actions, or act-tokens. Yet there are some corrections to be made to Goldman's notion of intentional action.⁶ Since it is the (second) best account I know of, showing the particular junctions where correction is needed is to take a step forward.

Goldman defines:

An *action-plan* consists of a *desire* (a *predominant* desire) to do some act *A'* and a set of *beliefs* (of greater or less certitude) to the effect that, if one were to perform basic act A_1 , this would generate (or be on the same level as) various other acts, including act *A'* (p. 56; my italics in '*action-plan*').

A basic act is essentially a bodily movement one can bring about just by willing it (p. 72). Then Goldman gives his crucial definition:

(G*.3) Suppose *S* has an action-plan which includes A_1, A_2, \dots, A_n , *A* is a basic act and $n \geq 1$. *S* wants to do *A*, and *S* believes (to some degree) of each of the acts A_1, \dots, A_n firstly, that it will either be generated by A_1 or be on the same level as A_1 , and secondly, that it will either generate A_n or be in the

same level as A_n . If this action-plan, IN A CERTAIN CHARACTERISTIC WAY, [2] causes S 's doing A_1 , then A_1 is intentional. And [3] if some of the other acts A_2, A_3, \dots, A_n are performed IN THE WAY CONCEIVED IN THE ACTION-PLAN, these acts are also intentional. [4] All other acts on the (actual) act-tree are NON-INTENTIONAL. (p. 57; the capitals and the labels express my added emphasis.)

Definition (G*.3) is admirably structuralist. Yet it suffers from some serious defects. To begin with, there are two connected deficiencies in the causal view on which it is erected. *First*, Goldman nowhere in his book explains the characteristic way in which the action-plan must cause the actions in it. Presumably he has in mind what lawyers call the “actuation” of the agent's action by his intention. But something special is added: a *characteristic* way of actuation, although it is restricted to the basic actions. For the other actions are caused in a way internal to the action-plan. It seems to me that he has a most important point here. Intentional action is caused in a certain special way given the agent's beliefs and desires. We will say something about that characteristic way below. We will see the causal involvement of practicalities, the neglected concept of action.

Second, Goldman leaves the elements in the causation of action shrouded in mystery. He says that the action plan causes the string of actions. The action plan is a complex of a desire and some beliefs about ends and means. This is *prima facie* wrong. Desires and beliefs are dispositional states. They can be causally involved, of course, as causal frames. A genuine cause is, however, the introduction of energy, it is an event. Goldman does note this later on and comes to speak of occurrent wants and occurrent beliefs. This is fine. Yet it is not a solution to the problem of how thinking causes action thereby becoming practical. What on earth is an occurrent want? An occurrent belief? I am not objecting to Goldman's introduction of occurrent entities as the causes of action. I am noting, with approval, that that move is in the right direction because it opens the problem.

We cannot simply speak of occurrent wants and beliefs and expect things to be clear. There is a serious travesty in speaking of occurrent wants and beliefs.

Dispositional states may consist of some internal occurrent states; but their dispositional character consists in a structural relationship pertaining to the succession of other states, namely: the occurrences that exercise or *rehearse* the dispositions in question. However, there need not be, and often there is not, any similarity between the internal occurrent state which a disposition is and the external occurrent states that exercise or manifest it. Just compare, for example, the internal occurrent state constituting a person's ability to swim, which state is in his body all the time, even when he is sleeping on a hard bed, with the motions of his muscles and the actual flow of energy from his brain to his limbs, when he is swimming. Furthermore, we understand the dispositional nature of a dispositional state in terms of the external occurrences that exercise or rehearse it. Thus, instead of leaving things obscure by casually speaking of occurrent belief and occurrent want, whatever they may be, we must know what sort of occurrences they are. Here is precisely where practitioners enter the stage.

A *third* serious defect of Goldman's explication of intentional action pertains to analysis (G*.3) itself. This analysis assigns too much significance to planned bodily movements. Goldman insists that:

- (i) a “complete action-plan ... i.e., a plan that includes a relevant *basic* act” is required for the intentionality of a goal action, and
- (ii) the plan must “*feature* the basic actions that *actually* generate” the goal act. (P. 60; my italics in 'feature' and 'actually').

Several things are wrong here.

A) Most action-plans that one formulates do *not* include bodily movements, which are Goldman's basic actions. These are somehow external to the conceived project we are proposing to carry out. We explain why in B). For instance, one plans to learn algebra or group theory and does not even think of the requisite bodily movements, e.g., how one will open the book, how one will lay out the pages to write the exercises, how one will tear out the latter pages, or fold them, when one moves to another page. Yet one learns group theory or algebra intentionally.

B) In fact the unity of a human project proceeds from above. The major objectives give sense to the project and all the subplans below have alternatives with which they are intersubstitutable – indeed, the lower the subplan the more substitutable it is. Clearly, every action is a bodily intervention. Hence, bodily movements, or nonmovements, are the entry points on every project, and are thus their basic components. Hence, *the bodily movements one produces in order to carry out a project are the most intersubstitutable elements of the project*. They are taken for granted and remain, thus, external to most projects or plans of action.

C) To act intentionally one does not need complete action-plans of the Goldmanian type. Indeed, complete action-plans in the sense of (G*.3) would destroy intentionality. That is, Goldmanian (G*.3) intentionality is inimical to the intentionality of actions that we use in daily experience.

To nail down these A)-C) let us consider some examples. For economy, let us exploit the tragedy of Francesca described above.

First example. Francesca's plan is to kill Romeo by shooting him with a gun she has just bought for that purpose. She grabs her gun from her purse and shoots Romeo. She is in a hurry and picks up her gun in such a way that she pulls the trigger with her right middle finger. She did not think of which finger to use. Thus, her plan of action did not include a basic action in Goldman's sense. Therefore, her action of shooting Romeo is not intentional on (G*.3).

Second example. Suppose that Francesca, a student of (G*.3), did put in her action-plan to pull the trigger with her right index finger. Even today (5 years later) Francesca thinks that she pulled the trigger with her right index finger, but she did actually fire her gun with her middle finger. In this case the agent has an action-plan with a requisite basic action; but it is the wrong actual action. Hence, her actions of killing Romeo, shooting him, and pulling the trigger – intentional in the normal sense – are not intentional by (G*.3).

Third example. Suppose that Francesca had one gun in each hand, and she had very carefully planned to shoot Romeo with both guns at the same time, missing him with the one in her left hand and hitting him in his heart with the gun in her right hand. However, she misses Romeo with her right-hand bullet, and her left-hand

bullet kills Romeo by destroying Romeo's brain. Obviously, she killed Romeo intentionally, but according to (G*.3) none of her actions can be said to be intentional.

As this case shows, bodily movements one can perform at will can be replaced with movements that are not even thought of. Indeed typically they seem to be performed in a reflex or habitual way. They are, when part of a complex project, intentional *only* by virtue of the project. They *cannot* bestow intentionality on actions causally generated by them, as Goldman's (G*.3) requires. Just the other way around: bodily movements gain any intentionality they may have precisely because they are the initial segments of processes whose terminal points, and some intermediate ones, are intentional.

This is, I believe, what ordinary language enshrines by not allowing us to describe Francesca's shooting of Romeo as having been done by Francesca with the intention of shooting him: this intention is not a terminal point in Francesca's plan of action. We must say, appropriately: She shot Romeo with *the* intention to kill him – not merely to shoot him.

Now, what we have noted about bodily movements (and basic acts) is also true of the intermediate stages in an action-plan, but *not* with the same degree of substitutability. We must, of course, investigate the principles that determine the boundaries within which those degrees fall. The central point is, however, absolutely clear: the integrity and unity of a plan of action comes from above, and the intentionality of the members of a plan of action also come from above, i.e., from the terminal point. Deviations from the terminal point are allowed, as in the case noted above in which Francesca shoots Romeo in his head having planned to shoot him in his chest. How much deviation is allowed? This is a question that has no definite answer. A careful consideration of examples, here skipped, reveals that the deviations from the original plan, without destruction of the intentionality of the actions whose characteristic states are brought about through those deviations, vary depending on the interests the community of the individuals involved have in those states. There is here an anthropocentric aspect of action, and of intentional action, that must be carefully studied. But we cannot go into them here⁷.

5. The main structure of intentional action: Data for any theory of intentional action

The preceding examination of Goldman's explication of intentional action is not complete⁸. It suffices to establish the following crucial points:

- (IA.1) Intentional action is, non-atomistically, an action that belongs to an intentional course of action.
- (IA.2) A course of action has a hierarchical structure, determined by a focal action, which may be called the *goal action*, or by several focal actions in the case of multiple-pronged project.
- (IA.3) A course of action that conforms within certain boundaries to a *project* (not exactly a Goldmanian action-plan) adopted by the agent is intentional.
- (IA.4) The precise boundaries of variations from the project allowed to a course of action for this to remain intentional are not precise: they depend on the moral or social gravity of the agents' purposes and actions, and also of how far from the focal action₁ of the project the variations are.
- (IA.5) No project is bound by the bodily movements of the agent, which are characteristically intersubstitutable – unless those bodily movements, as we shall see, are conceived only demonstratively.
- (IA.6) So-called basic actions and those episodes internal to the body of an agent that mediate between his volition and his bodily movements are in general only causally, not intentionally (i.e., not as part of the agent's projects), relevant to intentional courses of action.
- (IA.7) Exceptions to (IA.6) consist of projects whose goals are the production of bodily movement or even the episodes internal to the body from which episodes issue bodily movement.
- (IA.8) A project may include actions that lie beyond its goal actions in causal lines issuing from the agent's bodily movements that pass through the goal actions. They may be called *subsidiary consequences*.

6. First theoretical desideratum

We aim at the eventual formulation of a genuine theory of intentional action that conforms nicely and snugly to the eight criteria recorded above. The theory to be genuine must not consist of a mere string of definitions⁹. It must contain its share of definitions, of course, but it must be characterized by a constellation of laws, or principles or postulates, relating intentional action to other types of action, to volition, to belief, to circumstances, to value and deontic judgments, to causality and the transfer of energy¹⁰. In this study we shall concentrate on the flow of causality.

7. Causation and inference: conditional intentions

The holistic character of an intended course of action, we have seen, allows that a given volition covers, i.e., starts and monitors, a whole complex sequence of actions. This suggests that conditional intentions may be adopted by an agent and then put into realization by one over-all volition. If the volition that starts the realization of a conditional intention maintains its *Gestalt* character, then there would be no need for the insertion in the agent's agency of a special volition that brings about the realization of the conditioned intention. If no volition for the conditioned intention is to occur, then there are two alternative ways in which such intention could be realized. Either (a) the condition obtains, and by so obtaining it activates the agent's mechanisms involved in his attempting (or trying) to bring about the conditioned intention. Or (b) it is the agent's thinking that the condition obtains (whether it in fact obtains or not) that mobilizes the agent's actional powers toward his doing the action he conditionally intended to do. Since we are dealing with the agent's intentional action, what matters is how he *sees* the circumstances he is in. Hence, alternative (a) does not concern us here. We focus on alternative (b).

Let us consider an example. Suppose that:

- (1) At time *t* Bob Rosthal intends to do the following:
 - (i) if Jay Rosenberg visits him tonight, offer him his new French wine.

As the grammar of (1) indicates, the expression 'to do the following' being a scope indicator, what Rosthal intends is the whole of conditional (i). I am using the word 'intention' here to refer to the content of the state of intending. Thus, Rosthal's intention according to (1) is the conditional intention (i). This is as a whole conditional, but its condition is itself not an intention, but a circumstance. The primary text of intention (i) in (1) is of course Rosthal's first-person version:

(i*) If Jay Rosenberg visits me tonight, I will offer him my new French wine.

Rosthal is, therefore, in a dispositional state of intending, which itself is not conditional but whose *content* is a conditional intention. Now, to acquire the state of intending to do an action *A* is to create within one's body a network of routes for energy to travel, routes oriented toward the place where the relevant muscles and nerves are activated. The dispositional state of intending is *not* itself the activation of anything: it is merely that network of routes for energy travel. Thus, when Bob Rosthal made up his mind to offer Rosenberg his new French wine if he visited him, he literally *made up* his mind: he created channels for energy to go through. He arranged his mental structure.

What sort of energy can travel through those channels? Patently the energy that normally travels through nerves and muscles. Where is the energy required to move Rosthal's body toward his wine collection coming from? Here I am assuming a generally Humean account of causation to the effect that causation is fundamentally a relationship between events. Hence, we have to find an event that is to furnish the energy that can move Rosthal's limbs and muscles. The energy need not, of course, come from outside Rosthal's body. There must, however, be an event at least capable of mobilizing energy already available potentially in the body.

Clearly, for Rosthal to carry out his intention (i*) it is not sufficient that he offers his new French wine to a visitor who, unbeknownst to him, happens to be Jay Rosenberg. Indeed, it may even be the case that the appearance of the visitor in his house is what causes Rosthal to offer him his new French wine. Yet for the offering to be the intentional offering that fulfills the conditional intention (i*), Rosthal must

believe that his visitor is Jay Rosenberg. Some persons will even say that Rosthal *carried out* his conditional intention (i*) if he offers his new French wine to a visitor whom he mistakenly believes to be Jay Rosenberg. Such persons distinguish between Rosthal's carrying out intention (i*) and Rosthal's *doing* intentionally the action of offering Rosenberg his new French wine; for this doing to occur Rosenberg must be around.

In a literal sense it is not enough for Rosthal to carry out his conditional intention (i*) that he believes that Jay Rosenberg is visiting him. He must actually *think* that Rosenberg is *there* visiting him. The condition of conditional intention (i*) must be thought believably to obtain, and that episode of thinking must be causally involved in Rosthal's offering the wine to whom he (correctly) takes to be Jay Rosenberg. Some philosophers, however, even philosophers who hold a holistic view of intentional action, may want to argue that for Rosthal to carry out his conditional intention (i*) there must be another fact, or event, namely, Rosthal's going on to infer, from his conditional intention (i*), and his minor premise (that Rosenberg is there), the unconditioned consequent intention, thus:

- (i*) If Jay Rosenberg visits me tonight, I will offer him my new French wine.
- (2) Here is Rosenberg visiting me.
- (3) Hence: (ii) I am going to (will, shall) right now offer him my new French wine.

Clearly, Rosthal can make such an inference to an unconditioned intention. Then the *event* of inferring as follows, or the event of endorsing or adopting the conclusion, could certainly mobilize the required energy.

Undoubtedly, in some cases very likely Rosthal reasons “(i*), (2), therefore (ii).” Yet the issue is *not* whether he can make the inference, but whether he must make the inference in order to fulfill his initial conditioned intention (i*). Of course, we all agree that Rosthal can very well forget his initial conditional intention (i*), and upon seeing Rosenberg decide *anew* to offer him his most expensive French wine. If when he comes to intending conditional intention (i*) Rosthal creates in himself

a causal network of routes for energy to go through, it seems redundant to *require* that he recreates that network by rehearsing intention (i^*) as a major premise. Notwithstanding, sometimes that may be precisely what he has to do. A reasonable long time may have elapsed since he came to intend conditional intention (i^*), and he has been enjoying a very rich train of experiences, including the doing of many intentional actions of different sorts, and he may have, thus, blurred those routes. A process of clearing up the network of actual routes may be required, and an act of inferring “(i^*), (2); therefore, (ii)” may do just that.

Suppose, on the other hand, that Rosthal does not forget, i.e, his state of intending remains alive. Then, Rosthal made up his mind to do what conditional intention (i^*) formulates; he has not been enjoying experiences that touch the network of energy paths of which that state of intending consists. Thus conditional intention (i^*) is in the penumbra of, or just beneath, his consciousness. And while Rosthal is in that frame of practical mind, Rosenberg, with his characteristic flare, with no disguises or hats, comes in and greets Bob Rosthal with his unmistakable voice and style. Rosthal perceives Jay Rosenberg. Perceiving is an event, and one that brings in energy to Rosthal's brain. That event creates in Rosthal the dispositional state of believing that Rosenberg is visiting him, but that dispositional state is not involved in the mobilization of energy, although it is involved in the guidance of the energy to be mobilized. It is the event of Rosthal's perceiving Rosenberg, locating Rosenberg in his visual field through its demonstrative references, which mobilizes energy (whatever its ultimate source) and places it at the position where it will go through the clear routes that constitute the dispositional state of Rosthal's intending conditional intention (i^*).

It is purely an empirical matter whether or not the event of perceiving Rosenberg can mobilize enough energy within Rosthal's body for him to rise up, go to his wine cellar, choose from his French collection a rare Rothschild Lafitte, open it, and offer it to Jay Rosenberg. That is my point: it *is* an empirical matter. The empirical circumstances both within Rosthal's body and without his body can be such as to provide the clear channels for the requisite energy – mobilized by his perceptual thinking – to flow all the way. Rosthal's agency mechanisms may be well put

together with no relevant circuit broken; the immediate environment must be hospitable: there are no insurmountable obstacles in the way to the wine cellar, etc.

In brief, the connection between the volition to carry out conditioned action and the carrying out (or attempting to carry out) such an action is a causal one. It matches, and is grounded on, the connection between a major premise and a conclusion in a practical *modus ponens*.

The *Gestalt* character of intending explains how Bob Rosthal can proceed, having adopted conditional intention (i^*), from his thinking believingly (not merely entertaining i) that the antecedent of (i^*) obtains, directly, without the mediation of any inference and without a new volition to do the action mentioned in the consequent of (i^*), to the performance of such an action.

The above claim is based, partly, on the following general principle of implication that bridges the logical distance between practical and contemplative reason:

(Int.Bel*) If, at time t , X rehearses his intention to (A , if c), then if, at t , X rehearses his belief that c , then, at t , X intends to A .

This principle also bridges the distances between the dispositional state of intending to do a conditional action and the dispositional state of intending to do the action simpliciter (without the condition), through episodes of thinking. (Int. Bel*) is actually a weak principle: although it does not require that there be an occurrence of a thinking that rehearses the unconditioned intention, it requires a rehearsal of the conditional intention.

We also have a more general:

(Int.Bel**) If, during an interval of time d , X intends to (A , if c), then if time t is included in d and, at t , X has assertively in the penumbra or in the focus of his consciousness his intention to (A , if c) and, at t , X thinks believingly that c , then, at t , X intends to A .

These principles connect intending and believing along evident logical connections – by *modus ponens*! – that relate their contents. The best support for them lies in the holistic insight we gained in the preceding sections. Recall that because of the divisibility and the additivity of realized actions, any action can be considered as a sequence of smaller actions, or as a part of a larger action. We also noted that, regardless of how much divisible an action may be, intended actions must be considered as unitary wholes, whose divisions are junctures included in their very conception. We noted that it is utterly absurd to postulate a special volition for each possible segment that any division of a realized action can yield. Thus, any action, especially the more complex ones, can be treated as sequences of unitary actions. But here comes something truly significant: seldom is a unitary act sequence we intend to perform merely a juxtapositive succession of acts: acts have a *Gestalt*, as recorded in this principle:

(Act. Str.*) Seldom (if ever) an intended action, *A* which can be decomposed into the act sequence $A_1, \dots, A_i, \dots, A_n$ is such that each act A_i is intended by itself. Characteristically, the structure of the intention to *A* is of the form: intention to A_i & intention to (do A_2 , if A_1 is performed) & ... & do A_i , if A_{i-1} is performed. In short, every possible way of dividing an action that one intends to do yields a sequence of *conditionally* related actions in one's intending.

Recall that an agent does not think, or cannot even think, of each of the possible divisions of his actions into sequences. The agent has only a very schematic conception of the process that would obtain in the world were he to fulfill his intention. The present claim is that if the agent thinks of some possible division of his action, especially an action that he thinks of as an undivided unit, he would consider the latter acts of the sequence as intended conditionally under his realization of the preceding acts.

If the above sequentialization of any action whatever is correct, then the issue, whether or not a sequence of acts conceived as one unitary action – as in the case

of Ryle's clown – is matched by an isomorphic sequence of volitions, is precisely the issue whether or not a sequence of intended conditional acts is to be matched on a one-one basis with a sequence of volitions, or rehearsals of intending.

Perhaps the bridging implications recorded in (Int. Bel*) and (Int. Bel**) may be considered too rationalistic. Perhaps the reader may want to reject both principles and accept causal counterparts that include the additional condition that the agent *X* is rational. These weaker versions will certainly suffice for many purposes. Yet I believe that the stronger (Int. Bel*) and (Int. Bel**) hold also for irrational agents. These principles of implication determine neither which volitions the agent will have nor which actions he will perform. The principles merely tell that (dispositional) intending conditional intentions and occurrent believing that the conditions obtain imply the dispositional state of intending the unconditioned intention. The principles are non-empiricistic in that they allow that intentions can be adopted without having to go through consciousness. And this has to do with rationality only to the extent that it involves the rational power to think the believed and intended contents under consideration. But for the main causal thesis about intentional action I am proposing here, the weaker “rationalistic version” suffice. If the reader so desires he may read both (Int.Bel*) and (Int.Bel**) as having a built-in suitable antecedent, e.g. “*X* is at *t* rational with respect to his intentions.”

Now, Rosthal's example shows that in the case of intentions of the form *I to do A here now* we can have a kind of Aristotelian practical syllogism. We have an agent, Bob Rosthal, thinking practically of his conditional intention and thinking contemplatively or propositionally of the conditioning circumstance, and then he moves *directly* to action – without the need of further thoughts.

The example shows also that at most we can allow Aristotelian practical syllogisms whose conclusions are actions if the “premises” are at least mixed, containing at least one practical thought content. In this case such a practical thought content is the conditional intention. Furthermore, the example also shows that there can be genuine, non-Aristotelian practical syllogisms, whose conclusions are, not actions, but practical thought contents. One can derive an intention from another, and one can derive intentions from deontic judgments¹¹.

The case of Bob Rosthal is relatively simple. We have here an isolated conditional intention (or so it seems), and his act does not seem terribly complex. The nature of the phenomenon is perhaps clearer in the case of more complex projects. The analysis schema (Act.Str.*) provided above for the sequentialization of an action into conditional intentions should make the whole thing obvious. Clearly, once he perceives Rosenberg (as Rosenberg) – whether or not he derives intention (ii) from intention (i*) and (2) is immaterial – Rosthal will go to the wine cellar, etc. These actions through which he carries out his intention (ii), form if you wish, a train of acts, each of which Rosthal performs intentionally. He goes to his wine cellar intentionally, and intentionally walks down the stairs, and so each step he takes he takes intentionally. As long as there are no obstacles in the way, Rosthal performs – like Ryle's clown – a piece of habitual behavior. His total intention covers holistically the whole sequence of movements. But he will not have failed to fulfill his intention to take step 15, if his course of events is forced to abort at step 5 by the invincible obstacles. Step 15 enters in his plan, in his global intention, as one to be taken *if* step 14 is taken, and so on.

In general, habitual intentional action must be understood as a unitary pattern of acts that enter into the agent's intentionality in one piece, globally. The agent goes through each of the steps as if he had only contemplative consciousness of it. But that there is practical, volitional consciousness underlying it, undergirding the step to the previous steps, is manifested perspicuously when obstacles appear. An obstacle breaks the tranquillity of consciousness and demands on the agent's part a fresh new volition to continue his planned course of action, or to find a detour, or even to cancel the project. The greatest economy of intentional action requires that it must be possible for an agent – Rosthal, for example – to adopt sometimes a conditional intention and be moved to action by the thought that rehearses the belief that the condition obtains – without the mediation of an inference from the conditioned intention and the belief in the condition to the unconditioned consequent. Thinking takes time – as Plato taught us – and consciousness is too complex and too precious a commodity to be squandered away in unnecessary inferences. In an admirable universe – as I believe ours is, although not the best possible one, as Leibniz

proclaimed – neither thinking nor consciousness is squandered away. That is why we need consciousness and hard thinking when we are learning practices, acquiring habits, and overcoming obstacles. Thereafter, our behavior is habitual with great economy of consciousness and great efficiency. Yet habitual behavior may still be intentional and voluntary.

8. Brand's fundamental question of action theory

The preceding example also shows something relevant to Myles Brand's quest for the nature of the mental event that is the proximate cause of intentional action. Undoubtedly, a volition is a cause of action, at least of re-arrangements of inclinations to act and of the insertion of energy at some place available to the efferent nerves that will transfer energy to the appropriate parts of the agent's body that will be involved in the agent's doing of the willed action. But, as the example of Bob Rosthal shows, if our non-inferential view is correct, once a monitoring volition prepares the bodily paths for energy to go through, the actual insertion of energy through those paths need not be made by a volition, but by a thinking episode, often indeed an episode of perceptual thinking.

The preceding discussion shows how unnecessary it is to postulate two kinds of energy: one that goes with purely contemplative episodes, and another that belongs to practical thinking. Furthermore, one and the same type of energy is required to account for the transfer of energy to relevant parts of the body. We also see that propositional or contemplative thinking *can* be practical, if it occurs within a practical frame of mind.

9. Intending is a generalization of sorts of believing

The example shows three ways in which propositional thinking is involved in practical thinking. *First*, we have the fundamental fact that (1) implies (2) and (3) below:

- (1) At time t Rosthal intends to do the following: offer his new French wine to Rosenberg, if Rosenberg visits him tonight.

- (2) If at time t Rosthal believes that Rosenberg will visit him tonight, he intends then to offer him his new French wine.
- (3) If at time t Rosthal believes that Rosenberg will visit him tonight, he then both believes that he has some new French wine and intends to offer it to Rosenberg.

This implication suggests that intending is a kind of generalization of believing. Believed components can appear within an intending. *Second*, but not vice versa. The implication of (2) by (1) reveals that the circumstances or propositional components that can be transferred from intending into believing are not themselves pure contents of intending. A state of intending, whatever at this juncture its psychological reality may be, has to have its own proper contents. Note, for contrast, how (1) does *not* entail:

- (4) If Bob Rosthal intends Rosenberg to visit him tonight, he believes that he will offer him his new French wine.

It may be argued that (4) is senseless. What is the sense of locutions of the form “X intends Y to do A”? But we do not have to go to that extreme. The fact is that whatever the sense of (4) may be, it is not something (1) implies. Hence, there is a fundamental asymmetry between the components *Rosenberg visits him tonight* and *(Rosthal to) offer him (Rosenberg) his new French wine*. The former, a *circumstance*, expressed with a finite clause – an indicative clause – , is true or false, and is the appropriate content for believing: it naturally falls into believing, as the implicates (2) and (3) reveal. The other component, expressed with an infinitive clause, *must* remain in the scope of Bob Rosthal's intending. Grammar is a clue to the duality between circumstances and purely intended contents, i.e., intentions.

Thus, the second way in which propositional thinking is involved in practical thinking is that the mixed compounds of (propositions) circumstances and intentions are intentions. In mixed compounds the practical type is dominant, and the contemplative or propositional type is recessive. As the implication of (2) by (1)

shows, an understanding of the logic of intentions is crucial. But we will not discuss it here¹².

Third, as the Rosthal example also shows, the energy that causes an intentional action can be efficiently supplied by an episode of contemplative thinking – provided that such an episode occurs within a context structured or monitored by an appropriate comprehensive volition.

I have been pointing out features in which intending is a sort of generalization of believing, i.e., features in which intending seems to cover or include believing. Both are specific instances of endorsing a thought content. In believing we endorse a proposition (since believing is fundamentally *de dicto*, as the saying goes – even though the standard contrasts between *de re* and *de dicto* believing do not coincide, furthermore they are neither exclusive – the so-called *de dicto* has also a *de re* strand – nor exhaustive – what I have called quasi-indexical reference has been left out¹³). On the other hand, the endorsed contents of intending are, as we noted, intentions. There are more than fifteen different types of data that establish that an intention is not its corresponding proposition. For example, my intention to finish this paper today – note the infinitive – , which I express directly by a future-tense sentence, that is, my intention *I shall finish this paper today*, is not the corresponding first-person future-tense predictive proposition *I will finish this paper today*. Most of those data are discussed in *Thinking and Doing*; the main schedule of such data appears in Ch. 6, but the data are discussed in detail throughout the book.

The fact that intending and believing are both species of intending does *not* imply that there is an Aristotelian genus *endorsing*, which together with the respective specific difference constitutes believing and intending. Rather, both intending and believing are Johnsonian *determinates* of the determinable *endorsing*. This is a most important issue pointedly raised by Myles Brand in a beautiful paper¹⁴. The short of the story is that we must distinguish intending from believing in two dimensions: (i) by its content, and (ii) by its psychological reality. Just as we must not assimilate intentions to propositions, we must not assimilate the propensities, dispositions, and whatever occurrent structures constitute intending as a psychological reality with those that constitute believing as a psychological reality.

The sense in which intending is a generalization of sorts of believing, especially the third way, the causal one, does require that some of the structures and even propensities that may constitute a particular case of believing may intermesh with the structures and propensities that constitute a generalizing or covering intending.

10. Volition

A volition is an episode of thinking that introduces causation, or energy, into some appropriate place of the agent's body. It is an episode of practical thinking, more specifically, an episode of thinking an intention. Thus, we may say that it is a thinking episode that rehearses and exercises a state of intending – a dispositional state, as noted above – namely, the intending characterized by having as content or accusative the very same intention thought of in that thinking episode. In our human bodies, the appropriate place in which a volition introduces energy is a place in the brain where a bundle of efferent nerves receives the impulses which start the causal chain that will bring about the characteristic state of the action₂ the agent intends to do, provided, of course, that (a) the organism be well put together at least in that portion of it that is involved in the relevant bodily movement or position, and (b) the environment be hospitable, so that the causation of the bodily movement or position can be transferred all the way to the place where the characteristic state of the intended action can be caused to occur. Naturally, we must not prejudge how the bodies of other agents in other planets, or in this planet but in the future, can be wired or constructed. Thus, the efferent channels need not be nerves, and the energy that causes the characteristic state of the action need not be nervous in its initial stages. Here however we do not have to be concerned with these questions, which belong to the engineering of the agent's body.

11. The guiding element of an intention: The practical copula

What is of the greatest moment to attend to is that the intended thought content (i.e., the intention) *must* have some component, or feature, in some category or other that is neither subject nor predicate, that can guide the causation of practical thinking. That element must be systematically guiding. It must be present in the intention (=

intended content) so that the agent can know immediately, so to speak, by inspection of the contents of his mental events. It must be such that the thinking of the whole intention, with whatever circumstances it includes, can mobilize the agent's energy in the proper efferent channels. The appropriate channels are determined by the actions₁ contained in the primary intentions that are focal points of the global intention.

Evidently, the guiding element of the intention cannot be the first-person subject of the intention – because then the thinking of the subject with whatever action would make the complex an intention. Clearly, the agent must be capable of thinking some actions₁ as his own, without intending to do them, but instead as mere circumstances for his intentions. For example in somebody's intention: *If I encounter Mary at the Library today, I shall (am going to) ask her to come for dinner*, obviously the guiding element is not *I*. Certainly the antecedent is not a focal intended content, but is merely a conditioning circumstance. Likewise that systematic guiding element must not be a predicate – because then only such action as denoted by the predicate could be intended.

That guiding element must be a more formal or abstract component of an intention. It must be like (A) a copula, or (B) an abstract modifier of a predicate, or (C) a modifier of the copula, or (D) a modality of singular propositions (more explicitly, first-person present or future-tense propositions) that transforms the proposition in question into an intention. Views (A) and (B) are formally equivalent as is shown in *Thinking and Doing* Ch. 4. They can be regarded as surface views, which view (C) explains. This is shown in *Thinking and Doing*, Ch. 10. Moving from a practical copula (a special manifestation of it being the copula in intentions) to a practical modality on atomic propositions, i.e., moving from view (A) to view (D), may also provide an explanation of the copula. I have experimented with view (D).¹⁵ The preceding is a *transcendental argument* for at least the surface view (A). In this view the peculiar element in intentions (i.e. intended thought contents) that guides the mobilization of energy into the appropriate channels where intentional action has its immediate volitional cause, is a copula, to be distinguished from the propositional copula characteristic of propositional attitudes. The argument, in truly

Kantian spirit, merely seeks the necessary conditions of intentional action in terms of what the thought content of intentional volitional thinking must think to be causally efficacious. The argument connects the peculiarity of the causation of practical thinking involved in volitions with the contents of that thinking. Thus, the ARGUMENT pinpoints the difference that constitutes the determinate intendingly thinking and separates it from believingly thinking, both being determinates under the determinable endorsingly thinking. Patently, that difference is also the difference on the side of intending, between the determinates believing and intending, both being determinates under the determinable endorsing.

We need now the analogue of Kant's metaphysical deduction of the categories. For this we go to the form of complex intentions, where the differences in structure can be revealed. We have already discussed example (1) of Section 9, which shows the two copulae. The propositional copula is represented by the indicative form of the verbs, and the practical copula is represented by the infinitive form of the actional verb. For more on this see *Thinking and Doing*, Chs. 6, 7, and 10.

Thus, our transcendental argument illuminates the grammar of our ordinary sentences like (1)–(3) of Section 9.

12. The mind-body relationship

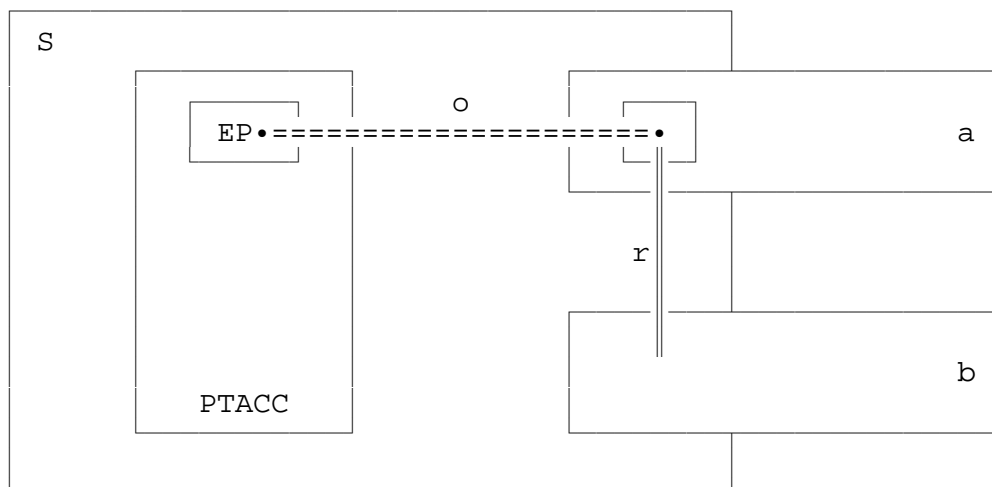
A volition introduces energy – or so it seems to our consciousness of doing something at will – at some place, the *Efferent Point*, in our brain, or, more generally, in our *Practical Thinking and Action Control Center*, where a process starts within the body of the agent that, if successful, will end up causing the characteristic state of the intended action. Here, we must stress, we are not doing the metaphysics of the mind-body relationship. It is not important for our present concern that a volition be a purely mental episode and that it introduce physical (chemical, electro-magnetic, kinetic, or whatever) energy into the Efferent Point, thus transforming mental or psychic energy into physical energy. Likewise, it is not important here that, on the contrary, there be no mental events that are not also physical events, so that there is no transformation of mental into physical energy. Similarly, there may be mental events that are not reducible to physical events,

volitions being perhaps some such events, that are caused by physical events, but are physically inefficacious. Or, for that matter, there may be a parallelism of events of the two types without inter-action.

Whatever the metaphysics of the mind-body relationship may be the fact is that in those situations in which the agent has the experience of a volition – whether by the volition itself, or by a physical event that causes the volition, or by a physical event that corresponds to the volition in the pre-established psycho-physical harmony – , energy is introduced in the appropriate Efferent Point. This is all that matters for us in the present context. (Of course, we want to understand the whole world, and we must take up the metaphysics of the mind-body relationship sooner or later.)

13. The internal causal diagrams

We consider here only bodies well put together (whether organisms or not, whether biological entities or not, whether the product of natural evolution of species or of human technology). We have, then, a volition to do some action impinging, by itself or through a corresponding physical event, in the Practical Thinking and Action Control Center at the appropriate Efferent Point. To build up momentum let us start with a simple volition, to wit, a volition to move a part of the body. For convenience let us consider an agent with the following simple design:



- *a* and *b* are movable limbs (or parts) of agent *S*.
- *PTACC* is *S*'s Practical Thinking and Action Control Center.
- *EP* is the Efferent Point at which the voluntary moving of limb *a* starts.
- *o* is the path of the energy within *S*'s body that the energy causing the motion of *a* will follow when *S* moves *a* voluntarily.
- Limb *b* has no paths of energy connecting it directly to *PTACC*; hence, there is no Efferent Point for *b*. But *b* is connected by mechanism *r* to *a*. If *a* moves, *b* will move.

For *S* to move limb *a* at will he must believe truly – whether such believing truly constitutes knowledge or not is a moot question – that his thinking the intention *I shall move a* (in whatever linguistic form he thinks of it is immaterial) at least *normally* (whatever this may mean) causes, is followed by, the motion of *a*. He may have that knowledge or true belief as an innate possession, or he may acquire it through different causes, including the special cause that may acquire it through different causes, including the special cause that may be called learning by experience. This cause consists in *S*'s having the thought *a will move*, or *I will move a*, or *I shall move a* and observing that in many of those cases *a* in fact moves. Again, learning is not our problem, and we can grant from the beginning that some Wittgensteinian point is met, namely, that those who learn to perform acts at will must somehow engage in social practices of some relevant sort.

Now, *S* has the power to move *a* at will. He does not have yet, let us suppose any idea about the constant conjunction between the motions of *a* and the motions of *b*. *S* does not need to know anything about the particular causal nature of the mechanism that goes from *EP* through *o* to the motion of *a*. But he *must* believe truly, and this is one of the most important claims I am putting forward here, that normally, i.e., *in normal circumstances*, his volition to move *a* is followed by (or causes) the motion of *a*. This belief is crucial. But *S* did not have any way of specifying what makes the circumstances normal, nor need he have any idea about what precisely the normal causation is. He must, nevertheless, believe, or take it for granted, that the circumstances he is in are normal.

In the example under consideration, normally, too, when S moves a he also moves b . Let us suppose that he acquires the true belief that in moving a he can move b . Then S can move b voluntarily through voluntarily moving a . And he can move a , hence, either as something that is an end in itself, or as a means for moving b .

Yet something else may obtain. Once S has a correct, though perhaps vague, belief about mechanism r connecting limbs a and b , he *can* intend to bring about the motion of b directly. All he needs is to take *de facto* – not, of course, *in cogitatione* – the efferent point EP for a as also the Efferent Point for b . He can then frame the intention *I shall move b* and rehearse the intending of this intention in an endorsing thinking of it. Now, if the causal diagram is clear, that thinking episode can be immediately volitional by inserting energy (of itself or appearing to do so as a reflection of the corresponding physical episode that can do so) at EP . Then the energy runs through path o (and through path r effecting the motion of limb b).

Undoubtedly, it is an empirical matter whether S will ever operate on the path $o + r$ as one internal causal path where his thinking intendingly of intentions can mobilize energy. What he needs is the belief that in *normal circumstances* he can volit the moving of b efficaciously. Again, he need not have any idea about the complexity of the path $o + r$. All he need is the true belief that there is a causal diagram internal to his body that connects his volition to move b with the motion of b .

14. Indexical reference in immediate volition

Let us return to the volitional content of the episode of thinking. We have already established that the content is not exactly a proposition, but a content that has a special formal element, a copula at first sight, that can monitor the introduction of energy at the appropriate Efferent Point. As noted above, the appropriateness of the Efferent Point has to do with the action one intends to perform. This suggests that the Practical Thinking and Action Control Center is also the Center where one has representations of propositions or states affairs. Let this be so. Our problem at the moment is that of one particular representation.

We continue with the story of *S*. How must *S* think of his limb *a* in his volition? What does '*a*' mean or stand for here, in the figure? Clearly, '*a*' is the name *we* gave *S*'s limb depicted to the right in our figure. But it need not be *S*'s own name to refer to his limb. Indeed, it is quite likely that *S* does not have a name for that limb, or for any part of his body. To be sure, *S* will have different ways of referring to limb *a* (in our current notation). Some ways of referring will be descriptions of the limb *a*, but perhaps he has in his idiolect a common noun to refer to limbs that perform in bodies like his a function similar to the one *a* performs. Suppose that that common noun is *kasha*, and suppose that *S* has only one *kasha*.

Suppose that one calescent afternoon *S* decides to move his *kasha*. In principle it could very well be the case that his thinking endorsingly *I shall move my kasha* (in a mixture of English and *S*'s language), where '*shall*', as before, expresses the intentional copula, can insert energy at *EP* and *S*'s limb *a*, i.e., his *kasha*, moves. In such a case, *S* could move his *kasha* at will without having a clear idea where his *kasha* is moving. If *S* had confused his *kasha* with limb *b* he would be moving *b*. If he had confused his *kasha* with another different limb for which there is a proprietary efferent point in *S*'s *PTACC*, then *S* would be moving this different limb.

The causal diagram above described is feasible. It connects the insertion of energy at *EP* with a description (property, predicate – take your choice here) of the bodily part to be moved. Like everything in life, it has its advantages and it has its disadvantages. One advantage is that it allows volitions to be effective, in normal circumstances, regardless of the location of the bodily parts, since it can operate by the conception of the function those parts play in the life of the agent. Thus, such a volitional mechanism would be of great value to an agent that does not move, and does not have to be concerned with the location of things. As we shall see below, the causal diagrams outside the body have a similar structure to those inside the body. Thus, an action that has a characteristic state far away from the agent's body occurs by the appropriate amount and quality of energy reaching the place where that characteristic state is to occur. Thus, a volition to bring about such a characteristic state must be able to send energy in the right direction so that this energy reaches

the right place. We must, therefore, as agents be concerned with the locations of objects that can be affected by our intentional actions.

Clearly, then, the advantage of a volitional agent whose volitions connect the insertion of energy at the appropriate place *EP* with attributes (properties, descriptions, or predicates) represented in contemplative part of the agent's Thinking Center is limited. The agent to be successful need not be concerned with the locations of the objects or persons he wants to affect with his actions. Naturally, such an agent, let us call him, *an attributionally volitional agent*, can succeed in realizing his intentions, if either (i) he limits himself to intentionally moving his bodily parts in accordance with certain functions they have, or (ii) he is capable of putting at the efferent point *EP* a *homing, attribute-seeking*, form of energy.

Obviously, human beings are neither limited to intentions to move their bodies or parts thereof nor can they muster attribute-seeking energy. Perhaps our successors in the chain of evolution – or, better, perhaps the technological products of our successors – will be attributionally volitional agents.

We have more limited powers. Our psycho-physical connection does *not* connect energy at the appropriate place *EP* with an attribute (or predicate). We decide to move our right index finger *right now*, and become confused and move our left index finger, or our right thumb. This, of course, can be because there is a crossing of the wires, so to speak, and we are attributionally volitional agents. But there is empirical evidence that that this is not so. One move *this* finger when one intends to move one's right index – and *this* finger is one's left index. In brief, one locates the object in one's current perceptual space and one volits to move the object perceptually so identified.

15. The identification of items in one's perceptual spaces is demonstrative

An immediate volition to move a part of one's body is, in our human condition, a volition to move *this*, or *that*. One can, indeed, intend to move the part of one's body that has property *F* and can carry out that intention. But at the moment of acting intentionally one has to identify the bodily part that is *F* as *this*, or *that*. Hence, in

the human agent's *PTACC* what inserts energy in the appropriate efferent point *EP* is a thought that includes demonstrative references to at least some of the objects and persons the agent intends to affect.

Patently what we have observed about the objects one intends to affect with one's actions applies *mutatis mutandis* to places and times. One identifies the places into which one has to throw energy immediately, given our human condition, as *here* or *there*. The same holds for time. One wills to do something *now*, and *here*.

In short, we, as human agents, are *indexically-volitional agents*. Our agency is governed by the following law:

(Int.Ind.) The content of a volition, which is a practically causal episode of practical thinking, is thoroughly indexical: It is a content structured by an I-Here-Now framework within which a This-That content is articulated alongside a practical (intentional) copula.

16. The causal transferability of indexical volition

Let us return to our agent *S* and his causal diagram depicted in the figure. According to the discussion in the preceding section, the immediate volition of *S*'s that moves his limb *a*, which he calls *kasha*, is an episode of intendingly thinking a volition of the form, *I'll move this* (referring to *a*, his *kasha*). If *S* has acquired the requisite beliefs to be able to move his limb *b* at will immediately, he has the power to mobilize energy into the efferent point *EP* that moves *a* and hence *b* by simply thinking *I'll move this* referring now to limb *b*.

Here we can see in a nutshell a powerful advantage of indexically-volitional agency. Indexical intentions allow a transfer of causation at will from an event that has a direct path to a point *EP* in the agent's *PTACC* to an event that has no such a direct path, but connects with one that does have such a direct path. And that transfer can take place without the possession of a mechanism that seeks attributes.

17. External causal diagrams and the intersubstitutability of subordinate action

So far we have, for the most part, stayed within the agent's body. Let us consider now actions that have characteristic states outside the agent's body. Let us take up the tragedy of Francesca di Verona again. Let's consider her action of firing the gun with which she shot Romeo. We have several crucial questions to ask:

- Did she fire the gun at will?
- Did she consider the gun trigger? How did she think of it?
- Did she think of her index finger with which she fired the gun?
- How many volitions did she have then?
- If she had one volition, what was the form of that volition?

There are many more questions one can ask. However, these suffice to help us gain additional understanding on voluntary and intentional action.

It seems safe to suppose that Francesca did have at least one volition when she killed Romeo. But it is not clear how many, and it is not clear that she had a volition for each of the actions she performed when she killed Romeo. Let us suppose, first, that she had a volition involving the firing of the gun. Then she thought something of the form *I'll fire now*. Of course, she need not have thought that. She may very well have thought *I'll kill him now*, where 'him' is an indexical. We will discuss this case later.

Francesca thought, then: *I'll fire now*. There are two sub-cases to consider. *First sub-case*: She had a little reflection to determine how to fire the gun and then a volition to press the trigger. *Second sub-case*: Francesca simply took all that for granted and merely volited the firing.

In the first sub-case we can pursue the matter further and have Francesca determine with which finger she will fire the gun. The whole proceeding has to end with her decision to *flex this (finger)*. If we start from here, then we are attributing to Francesca a sequence of volitions in accordance with the characteristic states of the actions we said that she performed. Patently, it is an empirical matter whether

she had one or many volitions. The crucial fact is the one we remarked above in our discussion of Goldman's explication of intentional action: the bodily movements and even many subordinate actions that lead to the doing of something intentionally are intersubstitutable. Thus, we may consider at once the second sub-case.

Francesca takes it for granted that she is holding a gun in her hand, that she is holding the gun trigger, that there is whole network of *normal* causal diagrams that connect her willing to fire and the firing of the gun. In this case, I submit, Francesca simply has one volition of the form *I'll fire now here* (referring to the place where her finger by flexing can fire the gun by receiving the energy issuing from her act of will), or of the form *I'll fire now by this movement* (referring now to the movement of her finger, whatever it may be, that is involved in the movement in question), or some alternative that thinks indexically of the relevant aspects of the situation.

To sum up, I am here proposing two theses:

(Int.Intersub.) The intersubstitutable subordinate actions in the intentionally carrying out of a certain project are actions that lie between the episode of volitional thinking and the action which in the volition itself is thought of indexically.

(Int.Norm.) The intersubstitutable subordinate actions are actions performed by the agent because they lie in:

- What must I everything considered do?
- What am I going to do?
- [the substance of the deliberation]
- Hence, I must everything considered do A.
- Therefore, I'll ('m going to) A.

(See *Thinking and Doing*, Chs. 2, and 10.)

Notes

An early version of this paper was presented at a symposium on action theory at the University of North Carolina at Greensboro in 1981.

¹ Aristotle, *De Anima* 434a16–22. Here he observes that the minor premise, which is indexical and of the *this is an act kind A and I am a person of type T*, really originates movement. He has in mind a deontic argument, whose major premise is of the form *Agents of type T must do actions of kind A*. For a detailed discussion of that text see Hector-Neri Castañeda, “Intentional Action, Conditional Intention, and Aristotelian Practical Syllogisms,” *Erkenntnis* 18 (1982): 239–260.

² Myles Brand, “The Fundamental Question in Action Theory,” *Noûs* 13 (1979): 131–151, p. 139.

³ Gilbert Ryle, *The Concept of Mind* (London: Hutchinson, 1949). Ryle thinks that a clown engaged in a skilful performance does not experience a volition just before each movement. And he is right. However that does not show that the clown does not execute several volitions during his performance, and, particularly, one over-all volition at the beginning.

⁴ Hector-Neri Castañeda, “The Conventional Aspects of Human Action, its Time, and its Place,” *Dialogue* 19 (1980): 436–460.

⁵ Alvin Goldman, *A Theory of Human Action* (Princeton: Princeton University Press, 1977), p. 12 and elsewhere.

⁶ See also Hector-Neri Castañeda, “Intentionality and Identity in Human Action and Philosophical Method,” *Noûs* 13 (1979): 235–260, pp. 236ff.

⁷ For the anthropocentrism of our very concept of action see the paper in *Dialogue* cited in note 4 above.

⁸ For further criticism of Goldman's analytic methodology and some of his main views see the paper in *Noûs* 79 cited above in note 6. For Goldman's response see his “Action, Causation, and Unity,” *Noûs* 13 (1979): 261–270.

⁹ For an assessment of the definitional methodology in philosophy see Hector-Neri Castañeda, *On Philosophical Method* (Bloomington, Indiana: *Noûs* Publications, 1980); see also the 1979 *Noûs* paper mentioned in note 6.

¹⁰ On causes and energy see Hector-Neri Castañeda, “Causes, Energy, and Constant Conjunctions,” in Peter Van Inwagen, ed., *Time and Cause* (Dordrecht: Reidel, 1980).

¹¹ For these matters and other issues pertaining to action theory, its logic, and its philosophy of mind, see Hector-Neri Castañeda, *Thinking and Doing* (Dordrecht: Reidel, 1975).

¹² See *Thinking and Doing*, Chs 4 and 6.

¹³ See my “Reference, Reality, and Perceptual Fields,” Presidential Address, in *Proceedings and Addresses of the American Philosophical Association* 53 (1980): 763–823.

¹⁴ Myles Brand, “Intending and Believing,” in James Tomberlin, ed., *Agent, Language, and the Structure of the World* (Indianapolis: Hackett Publishing Co., 1982): 171–193. See also “Reply to Myles Brand: Intentions, Properties, and Propositions,” *Ibid.*: 411–417. Brand included a rejoinder in his *Intending and Acting* (Cambridge, Massachusetts: MIT Press, Bradford Books, 1984): 85–100, 272–273. I am not persuaded. We disagree on the nature of determinable and determinate properties, on the double difference I have insisted on between intending and believing, and on the nature of intentions. Brand's view of intending is very limited; he declares: “I do not discuss conditional intentions” (p. xi), which intentions provide both my major data and one of my chief theoretical targets. Conditional intentions are particularly important because they are mixed: their conditions are propositions and their conditions are practitions. See further *Thinking and Doing*, Ch. 6. Concerning types of family of properties see, Hector-Neri Castañeda, “Negations, Imperatives, Colors, Indexical Properties, Non-existence, and Russell's Paradox,” in D.F. Austin, ed., *Philosophical Analysis* (Dordrecht: Reidel, 1988): 210–234, and “Indexical Thinking and Indexical Guises (Reply to Harald Pilot),” in Klaus Jacobi and Helmut Pape, eds., *Das Denken und die Struktur der Welt – Thinking and the Structure of the World* (Berlin: De Gruyter, 1990): 299–307.

¹⁵ See Hector-Neri Castañeda, “Ought, Time, and Deontic Paradoxes,” *The Journal of Philosophy* 74 (1977): 775–791.

PHILOSOPHICAL PUBLICATIONS, 1986–1990

This bibliography supplements the bibliography in the 1986 Tomberlin (ed.) volume (see below). It contains Castañeda's publications up to November, 1990.

Books

- 1 *Sul metodo in filosofia*, (Trento: Luigi Reverdito Editore, expanded and translated by Roberto Paoli, 1989).
- 2 *Sprache und Erfahrung: Texte zu einer neuen Ontologie* (Frankfurt: Suhrkamp, translated by Helmut Pape, 1982).
- 3 *Thinking, Language, and Experience* (Minneapolis: University of Minnesota Press, 1989).

Main contributor in the following books in his honor

- 4 James E. Tomberlin, ed., *Hector-Neri Castañeda* (Dordrecht: D. Reidel, Profiles Vol. 6, 1986) – “Self-profile” and 8 replies [TOMBERLIN].
- 5 Klaus Jacobi and Helmut Pape, eds., *Das Denken und die Struktur der Welt: Castañedas epistemische Ontologie in Darstellung und Kritik* (Berlin: De Gruyter, 1990) – 4 autonomous essays and 17 replies [JACOBI-PAPE].

Essays

- 6 “Intellectual Autobiography: De Re” in TOMBERLIN: pp. 3–77.
- 7 “Self-profile: De Dicto” (including “Self, Thinking, and Reality”), TOMBERLIN: 77–137.
- 8 “Reply to Jay Rosenberg,” TOMBERLIN: 333–340.
- 9 “Reply to David Woodruff Smith,” TOMBERLIN: 340–348.

- 10 "Reply to Esa Saarinen," TOMBERLIN: 348–350.
- 11 "Reply to Stephen Boer and William Lycan," TOMBERLIN: 351–370.
- 12 "Reply to Frank McGuinness," TOMBERLIN: 370–373.
- 13 "Reply to James Tomberlin," TOMBERLIN: 373–375.
- 14 "Reply to David Wong," TOMBERLIN: 375–379.
- 15 "Reply to Bruce Aune," TOMBERLIN: 379–382.
- 16 "Reply to Jeffrey Sicha," TOMBERLIN: 382–391.
- 17 "Underdevelopment and Resources: The need of a Structural and Global Approach," Werner Leinfellner and Franz N. Wuketits, eds., *The Tasks of Contemporary Philosophy* (Vienna, Hölder-Pichler-Tempsky, 1986): 227–231.
- 18 "Aspectual Actions and Donald Davidson's Theory of Events," Ernest Lepore and Brian McLaughlin, eds., *Actions and Events* (Oxford: Blackwell's, 1985): 294–310.
- 19 "Obligations, Aspectual Actions, and Circumstances," *Philosophical Papers* 15 (1986): 155–170.
- 20 "Practical Reasons, Reasons for Doing and Intentional Action (The Thinking of Doing and the Doing of Thinking)," *Theoria II* 2 (1986–87): 69–96.
- 21 "The Self and the I-Guises, Empirical and Transcendental," in Konrad Cramer, Hans Friedrich Fulda, Rolf-Peter Horstmann, und Ulrich Pothast, eds. *Theorie der Subjektivität* (Frankfurt a. M.: Suhrkamp, 1987): 105–140.
- 22 "Self-consciousness, Demonstrative Reference, and the Self-ascription View of Believing," in James E. Tomberlin, ed., *Philosophical Perspectives: I. Metaphysics* (Atascadero, California: Ridgeview Publishing Co., 1987): 405–454.
- 23 "Existence, Object, and Reference," *Grazer Philosophischen Studien* 25/26 (1985/86): 3–64.
- 24 "Direct Reference, the Semantics of Thinking, and Guise Theory (Constructive Reflections on David Kaplan's Theory of Indexical Reference)," Joseph Almog, J. Perry, and H. Wettstein. eds., *Themes from David Kaplan* (Oxford: University Press, 1989): 105–144.

- 25 "Knowledge and Epistemic Obligation," James E. Tomberlin, ed., *Philosophical Perspectives: II. Epistemology* (Atascadero, California: Ridgeview Publishing Co., 1987): 211–233.
- 26 "Persons, Egos, and I's: Their Sameness Relations" in M. Spitzer, F. A. Uehlein, and G. Oepen, eds., *Psychopathology and Philosophy* (Berlin Heidelberg: 1988): 210–234.
- 27 "Metaphysical Internalism, Selves, and the Holistic Indivisible Noumenon," *Midwest Studies in Philosophy* XII (1988): 129–144.
- 28 "Semantic Holism Without Semantic Socialism," *Midwest Studies in Philosophy* XIV (1989): 101–126.
- 29 "The Semantics of Thinking, Dia-philosophical Pluralism, and Guise Theory," *Metaphilosophy* (1988): 79–104.
- 30 "The Basic Logic for the Interpretation of Legal Texts," in Charles Walter, ed., *Computer Power and Legal Language* (Westport, Greenwood Press, 1988): 167–181.
- 31 "Philosophy as an Institution and as a Worldview," in Avner Cohen and Marcelo Dascal, eds., *The Institution of Philosophy: A Discipline in Crisis?* (Chicago: Open Court, 1989): 35–60.
- 32 "The Reflexivity of Self-Consciousness," in Christopher Hill, ed., *Philosophical Topics* 17 (1989): 27–58.
- 33 "Supervenient Properties and the Hierarchy of Emergent Individuals," Hemult Pape, ed., 1987 Proceedings of the German Semiotic Society: Peirce Section (forthcoming).
- 34 "Objects, Identity, and Sameness," *Topoi* Supplement 4 *The Object and its Identity* (1989): 31–64.
- 35 "Paradoxes of Moral Reparation: Deontic Foci and Circumstances," *Philosophical Studies* 57 (1989): 1–21.
- 36 "Moral Obligation, Circumstances, and Deontic Foci (A Rejoinder to Fred Feldman)," *Philosophical Studies* 57 (1989): 157–174.
- 37 "The Content of Legal Speech Acts and Legal Deontic Logic," *Rechtsphilosophie en Rechtstheorie* 18 (1989): 108–129.

- 38 “The Multiple Faces of Knowing: The Hierarchies of Epistemic Species,” in J. W. Bender, ed., *The Current State of the Coherence Theory* (Dordrecht, Boston, London: Kluwer Academic Publishers, 1989): 231–241.
- 39 “Perception: Demonstrative Subjective Aspects and Non-Communitarian Sameness,” in Wolfgang Gombocz, Heiner Rutte, and Werner Sauwer, eds., *Traditionen und Perspektiven der analytischen Philosophie* (Verlag Holder-Pitchler-Tempsky, 1989): 284–296.
- 40 “Uncertainty in Morals and Politics,” in George von Fürstenberg, ed., *Acting under Uncertainty: Multidisciplinary Conceptions* (Boston, Dordrecht, London: Kluwer, 1989): 232–264.
- 41 “Die Ontologie hat ihre Bedeutung nie verloren: Ein Gespräch mit Hector-Neri Castañeda,” [an interview with H-N. C. by Gerhard Lob and Michael F. Loebert], in Dieter Birnbacher, ed., *Information Philosophie* 1 (1989): 43–50.
- 42 “Lógica deóntica, sus paradojas y una solución uniforme” in Francisco Miró Quezada, *Lógica en Latinoamérica* (forthcoming).
- 43 “Leibniz's Complete Propositional Logic,” *Topoi* (special issue on Leibniz's logic edited by Massimo Mugnai), Vol. 9, No. 1 (1990): 15–28.
- 44 “Language, Thinking, and Reality,” Sachowiecz, ed., *Pragmatics*, Vol. IV (forthcoming).
- 45 “Comment on Prof. Weimar's View of the Moral Sciences” in Gerard Radnitzky, ed. *Centripetal Forces in the Sciences* (New York: Pergamon Press, 1989).
- 46 “Practical Thinking, Reasons for Doing, and Intentional Action: The Thinking of Doing and the Doing of Thinking,” in James E. Tomberlin, *Philosophical Perspectives*, Vol. IV (1990): 273–308.
- 47 “The Role of Apperception in Kant's Transcendental Deduction of the Categories,” *Noûs* 24 (1990): 147–157.
- 48 “Proper Names, Singular Reference, and Baptims” (Reply to Helmut Pape), JACOBI-PAPE, pp. 195–201.
- 49 “Proper Names, Personal Pronouns, and Free Variables” (Reply to Guido Küng) JACOBI-PAPE, pp. 207–210.

- 50 “Proper Names, Variables, and Reference” (Reply to Hans Ineichen) JACOBI-PAPE, p. 218–229.
- 51 “Proper Names, Semantics, Guises, and Doxastic Referents” (Reply to Paolo Leonardi) JACOBI-PAPE, pp. 244–258.
- 52 “Fiction, Perception, and Elliptical Speech” (Reply to Wolfgang Kuenne) JACOBI-PAPE, pp. 268–284.
- 53 “Indexical Thinking and Indexical Guises” (Reply to Harald Pilot) JACOBI-PAPE, pp. 299–307.
- 54 “Demonstrative Reference, and Kant's Copernican Revolution” (Reply to Gerold Prauss) JACOBI-PAPE, pp. 312–319.
- 55 “Perception, Subjectivity, and Thinking” (Reply to Richard Schantz) JACOBI-PAPE, pp. 329–338.
- 56 “Perceptual Experience, Individual Guises, Physical Objects, and the World” (Reply to Anton Friedrich Koch) JACOBI-PAPE, pp. 348–362.
- 57 “Doxastic Reference, Thinking Reference, and Demonstrative Individuals” (Reply to Frank Doering) JACOBI-PAPE, p. 372–384.
- 58 “Thinking Reference, Gestures, and Communication” (Reply to Rainer Enskat) JACOBI-PAPE, pp. 401–407.
- 59 “Individual Guises, Spatiality, and Reality” (Reply to Guido Küng) JACOBI-PAPE, pp. 416–424.
- 60 “Intentionality, Perception, and Guise Theory” (Reply to Heinz-Dieter Heckmann) JACOBI-PAPE, pp. 438–448.
- 61 “Individuals, Reference, and Existence” (Reply to Tomis Kapitan) JACOBI-PAPE, pp. 459–472.
- 62 “Ontological Perspectivity and Guise Theory” (Reply to Friedrich Rapp) JACOBI-PAPE, pp. 477–481.
- 63 “Forms of Predication” (Reply to Hans Burkhardt and Carlos Dufour) JACOBI-PAPE, pp. 491–494.
- 64 “Individuals, Idealism, and the Realism of Thinking” (Reply to Klaus Jacobi) JACOBI-PAPE, pp. 523–546.

- 65 "Self-Consciousness, I-Structures, and Physiology," in Manfred Spitzer, ed., *Philosophy and Psychopathology* (New York, Springer-Verlag, 1990): 118–145.
- 66 "Leibniz's Views on Contingent Truth in the Late 1680's," Wolfgang Lenzen, ed. [Festschrift for Heinrich Schepers] 1990 (forthcoming).
- 67 "Morals, Metaphysics of," in Barry Smith, ed., *Handbook of Metaphysics and Ontology* (Munich, Philosophia Verlag GmbH, forthcoming).
- 68 "Perceptual Individuals, Demonstrative Reference, and the Semiotics of Perceptual Judgments," International Peirce Congress, Cambridge, Mass., Harvard University, Sep. 5–10, 1989.
- 69 "Objects, Existence, and Reference: A Prolegomenon to Guise Theory," Donald Stewart, ed., *Entities and Individuation* (Lewiston, Edwin Mellon Press, 1989): 39–87.
- 70 "Indexicality: The Transparent Subjective Mechanism for Encountering a World," *Noûs*, Vol. 24, No. 5, (1990 forthcoming).
- 71 "Leibniz's View of Existence in 1689," *Stoicheia*, (1990): 5–17.
- 72 "Elección, Constitución, y Democracia," *Guatemalan Judiciary Review*, (forthcoming).