

# *The Reflexivity of Self-Consciousness: Sameness/Identity, Data for Artificial Intelligence*

Hector-Neri Castañeda  
*Indiana University*

Phenomeno-logic is the study of the logic of phenomena. Phenomena are what appears in experience. That is, the phenomeno-logic of X is the logic of the consciousness of X. . . . The phenomeno-logic of self-consciousness studies the structure of the consciousness of the appearance of self. To carry out such a study is to be conscious of the form of self-consciousness as this is experienced. . . . One major tool of phenomeno-logic is phenomeno-logical linguistics. It investigates how the contents of a type of experience become encoded in the language used to live that experience.

—Tor Daschein, *The Reflexivity of Self-Consciousness*

## I

### PROBLEM, ASSUMPTIONS, AND MAJOR THESES

Consciousness (in the primary sense of this word) is an *occurrent* condition. It is a general condition, as when, for example, we re-gain consciousness by coming out of surgical sedation or a coma. Crucially, this condition permeates, and exists in, particular episodes of being conscious of this or that. There is, however, a dispositional sense of 'consciousness'. For instance, Marxists speak of the plutocrats' class consciousness and of the so-

cial consciousness of the labor class. This consciousness is not (occurrent) consciousness, but a complicated network of attitudes and propensities to act in special ways, tolerate certain types of actions, promote certain species of behavior, experience feelings of certain kinds, have beliefs of certain types, and argue in appropriate ways. These attitudes and propensities exist even though they are not manifested, but when manifested they will naturally deploy themselves in episodes of consciousness of the pertinent contents.

Here we consider occurrent consciousness only. More precisely, we focus on episodes of *self*-consciousness. These are episodes in which ONE is conscious of ONEself *qua oneself*. Our topic is the reflexivity of self-consciousness. The reflexivity in ONE referring to ONEself as *oneself* is twofold. There is the external reflexivity of ONE referring to ONEself, and the internal reflexivity of ONE referring to something, whatever it may be, as *oneself*. We must take both reflexivities into account. The internal reflexivity is the peculiar core of *self*-consciousness.

*Self*-consciousness is executed in episodes of thinking about oneself *qua oneself*. The contents thought in such episodes are expressed in utterances of sentences containing (at least apparently) singular-referring uses of the first-person pronoun. For this reason we may more accurately speak of *I*-consciousness. A thinking episode is of course not an event of uttering. It is embodied in—indeed, in some *appropriate* sense of ‘sameness’, a thinking episode is the same as—an event or process in the thinker’s brain, or whatever the thinking box of the thinker may be. Here we neither endorse nor oppose the reductionist physicalistic thesis that mental events are just second-order causal properties of physical events. We must, however, insist that mental events and mental dispositions are in the world and are part of the causal order. We must accept causal equations of the mental and the physical, and raise the fundamental issue of the nature of the sameness involved in such equations. Here, however, we do not enter into this issue.<sup>1</sup> One thing is clear to me: the sameness in question cannot be conceptual or analytic equivalence, much less literal (self)identity: hence, such equations cannot provide a reduction of the mental, more specifically, of consciousness, to the physical. Yet they may secure the causal dependence entrenched in the hierarchical emergence of mental states and particulars on and off physical states. In fact, consciousness seems to be an irreducible emergent. This is, however, not the occasion to tackle this topic.<sup>2</sup> In any case, reductionist functionalists are, or should be, firmly concerned with the reflexivity of self-consciousness. The better we understand what it appears to be, the more detailed and secure our reductionist programs can be. Likewise, Artificial Intelligence, whether practiced with a reductionist bent of mind or not, has a vested interest in the reflexivity of self-consciousness. Clearly, the production of facsimiles of human behav-

ior or of mental states and activities needs only the causal dependence of the mental on the physical. Self-consciousness is the apex of mentality.

Here we are not concerned with development, but with the structures of self-consciousness. Our subjects are mature, fully competent thinkers who enjoy or suffer full episodes of *self-consciousness*. Such subjects have acquired the requisite abilities to think thoughts of the form *I am such and such*. We assume that they have mastered a natural language by means of which they relate to others, with whom they engage in social activities and practices. In brief, our mature thinker has a natural language that she uses both as a means of communication and as a means of thinking. Thus, there is as good a match as is feasible between our thinker's brain events embodying her thinking episodes and some linguistic utterances available to her that express the contents of such episodes. That is, her sentence tokens in her varying contexts of thought and speech can formulate the internal contents constitutive of those thinking episodes. We further assume for convenience that our mature thinker is, at least in the cases we discuss, *candid*: her sentence tokens reflect, within the normal limits and subject to applicable standard contextual constraints, what she is thinking at the time of utterance.<sup>3</sup> The assumption of speaker's candidness is subordinate to the chief assumption that language is spoken. To be heard, or overheard, language must be spoken. Thus, though we are interested in the hearer's perspective, we are deeply concerned with the speaker's point of view. This involves a needed departure from the standard philosophizing carried out in English. Surely to investigate *self-consciousness* is to attempt to get into the thinker's most intimate standpoint. Nonetheless, assuming a hearer, we initially pose the problem as the exegesis of attributions of self-consciousness. This is the problem of ascertaining what a speaker thinks and what a thinker thinks when the former attributes self-consciousness to the latter.

Armed with the preceding methodological assumptions we can discuss the thought contents of episodes of *self-consciousness* as they manifest themselves in the pragmatico-semantic content of tokens of first-person sentences. We need not worry about the neurophysiology or the engineering of the thinker's processes embodying such episodes. For our task here it does not matter whether the sequential order constitutive of such processes is compositionally syntactic or connectionist, or whatever.<sup>4</sup>

At the risk of provoking hostile opposition, as well as greater expectations than can be fulfilled here, I want to deploy the general view of *self-consciousness* urged by the examination of central data. It is constituted, among others, by the following theses. *First*, contrary to a widely held Fichtean tradition, not all consciousness is *self-consciousness*. *Second*, the different degrees of consciousness are essentially cumulative and hierarchical: more advanced degrees not only presuppose but include suffusively the

lower degrees. (This does not show that brain events embody thinking or consciousness compositionally. It does show, however, that the surface operations of the brain, those that ultimately deliver consciousness, have a general compositional syntactic structure.) Thus, *self-consciousness* is built up on layers of *self-less* consciousness which remain as an internal basis for reference to oneself as *I*. *Third*, consequently, the unity of an episode of consciousness is not explicable by virtue of that consciousness belonging to a self, or, better, an *I*.<sup>5</sup> On the contrary, it has to be accounted for in its own terms. Besides, if an episode of consciousness is internally owned by an *I*, then the unity of that owned consciousness is, rather, an element in the constitution of that ownership, and an internal presupposition of that *I*. *Fourth*, the hierarchy of content of full *self-consciousness* is twofold. On the one hand, it includes a structural hierarchy of formal contrasts—a structure of *conceptual negativities*. On the other hand, those contrasts have to be realized in specific contents—a network of *empirical positivities*. These are identifications of what one is, and raise a question about the sense of the requisite identity or sameness.

*Fifth*, the vital role of an episode of *self-consciousness* is exhausted in current experience: it is organizing and monitorial. Here lies one part of the solution to the biographical paradox of *self-consciousness*. Consciousness is an expensive commodity worth sparing. To live is fully to immerse oneself in action, that is, in full obliviousness of oneself, developing techniques of habitual and immediate, but adequate response to problems. Conversely, as Plato remarked,<sup>6</sup> the best life is an *examined*, a fully consciously examined life; yet an *examining* life is not the best form of living—if living it is. *Sixth*, the experiential function of first-person reference is of one piece with the experiential role of all other indexical reference. *Seventh*, THE MECHANISMS OF INDEXICAL REFERENCE CONSTITUTE THE STRUCTURE OF SUBJECTIVITY. The strict referents of acts of indexical referring are fleeting subjective particulars that serve to harpoon the external referents being experienced. In the case of perception, we aim at harpooning, and we typically harpoon, (perceived) physical entities; in the case of first-person reference, we harpoon thinking persons with their bodies. *Eighth*, hence, indexical particulars, which are ephemeral and subjective, have a twofold epistemological primacy. On one side, indexical particulars enjoy a basic intrinsic incorrigibility: they exist as and inasmuch as they are experienced. On the other side, since their role of harpooning external reality is their *raison d'être*, their subjective incorrigibility typically and normally transfers to the physical realities they harpoon as the latter's immediacy of presentation in experience. The ontology of indexical particulars is wholly exhausted in their epistemological roles. *Ninth*, first-person reference is true to its indexical character. The *I*'s,

strictly referred to in episodes of *self*-consciousness, are evanescent particulars whose function is to harpoon, through reflexive thinking, persons as subjects of experience, whatever they may otherwise be. *Tenth*, one peculiar feature of the *I*'s is their effective transfer of their incorrigibility to the reality they harpoon for the experiencer. (This is the Cartesian truth of the fundamental certainty of the *cogito*.) *Eleventh*, there are no lasting *I*'s. This raises an interesting problem about first-person past- and future-thoughts and statements. Of course, underlying each evanescent *I* there *may*, logically, be an enduring physical entity, or a sequence of such entities, or even an eternal soul. However, the ontology of the *I* grounds by itself neither an empirical theory of a particular embodiment of consciousness nor a metaphysical theory of the soul.

## II

### THE REFLEXIVITY AND THE SELF-REFERENCE OF SELF-CONSCIOUSNESS

Above we have introduced our topic with the formula:

(I\*) In an episode of *self*-consciousness ONE refers to (thinks of) ONEself as (*qua*) *oneself*.

We must proceed to a patient exegesis of this formula. To begin with, the meaning of the prefix 'self-' in 'self-consciousness' captures the interplay of the pronoun sequence 'ONE . . . ONEself . . . *oneself*' in (I\*). It also absorbs the roles of the suffix '-self' in (I\*). The sense of the suffix '-self' 'ONEself' is very different from its sense in '*oneself*'. The former expresses an attributed external reflexivity; the latter depicts a lived internal reflexivity.

The first occurrence of '-self' in (I\*) is merely a logico-syntactic device signaling that the variable 'ONEself' has both the same value and the same substitution instances as, is co-referring with, really is another token of the same variable as, the preceding 'ONE'. That '-self' is just like the one in "John, unawares, cut himself," attached here to 'him' to signal that 'himself' is the same singular term 'John'. In brief, the tandem 'ONE . . . ONEself . . .' exhibits that formula (I\*) is of the form:

(I\*.f1) X refers to X as . . .

In its turn, form (I\*.f1) is a specialization of the form:

(GR) X refers to Y as Z.

Thus, the external reflexivity of 'ONE refers to ONEself' is merely the reflexivity of any relation *R* whatever that has in special cases the same

argument (relatum) in two (or more) positions: R(. . . x . . . . x . . .). This external reflexivity appears in trivial comparative relations, e.g., being as old (tall, thin, beautiful, etc.) as itself. There are however non-trivial external reflexivities. A wire under pressure may come to touch itself; a very hungry tiger that eats his own arm eats of himself, wounds himself, feeds himself with himself. This reflexivity needs not be mental; nevertheless, it is the most basic ingredient of *self*-consciousness.

Let us examine the *as*-moment in the general formula (GR) *X refers to Y as Z*. This is a requisite background for our study of the moment *as oneself* in self-consciousness.

An act of referring is a real event in the world. It is, therefore, a part of the causal order. It involves a relation which can have the self-same entity in the positions of agent and accusative. Hence, there is a causal relation that includes an external dovetailing of reference on to its source. (This is a wonderful topic for another occasion.)

An act of referring is a mental event. As such it has an internal content. This is what the sub-form 'as Z' of (GR) alludes to. Content Z is what the thinker refers to, or thinks of. It is what thinker X is allegedly *presented with*. The physical-physiological process of thinking something about Z, say, that (. . . Z . . .), is undoubtedly thoroughly computational. It may include a multitude of computations of brain states, terminating in a complex B of brain events that contains a representation of that (. . . Z . . .) and of Z. That is, either there is an isomorphism between some systematic parsing of that (. . . Z . . .) and B, or there is a mechanism for constructing such an isomorphism, or there is a general causal function that assigns (. . . Z . . .) to B. However, for X to think that (. . . Z . . .) the occurrence of that representational event is not enough. That representational event must deliver a *presentation to the thinker* of what he is thinking of. This is of the greatest importance. The representations of what the thinker refers to may be stored in different ways in his brain. This storage is not thinking. Those representations may move from storage to occurrence, yet this transition may not suffice. This is so even if the representation were a perfect replica or image of the object thought of, and the thinking occurrence of the image consisted in its being a faithful replication of the stored image in the thinking box, screen, or whatever, in the brain. Certainly, to embody an act of referring (or purporting to refer), the occurring representation of the referent must occur in the proper location required by the engineering of the brain. It must, further, occur in the proper *presentational* way: It must yield the molar state of the possessor of that brain being presented with Z and with (. . . Z . . .). Here is the emergent phenomenon of *consciousness*, accruing to a thinking whole. The purely mental and crucial sense of '[refers to, thinks of] as (*qua*) Z' is governed by this fundamental principle of the embodiment of thinking:

(Th.PR) In the general non-existential sense of ‘refers to’ and ‘thinks of,’ what the referrer is presented to is the *self-same* item as what he refers to (thinks of).

In terminology I do not like, the locution ‘as Z’ expresses a *de dicto* attribution of a mental act. As always, *de dicto* attribution is here an attribution by depiction of the internal, presentational content of such an act. Suppose I declare: “John believes that Leibniz was a lawyer.” Evidently, my subordinate clause discloses to the hearer, not John’s words—let alone John’s brain representations—but a sort of replica or a picture of John’s belief content.

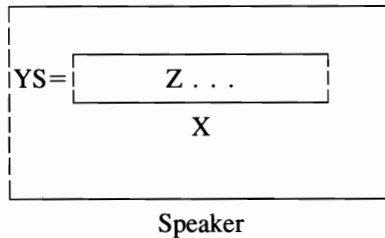
The formula (GR) *X refers to Y as Z* is of a mixed nature. The component *to Y* is, as we said, external. It is, in one sense of a companion terminology I also dislike, *de re*. The point often made with this expression is that what a substituent of ‘Y’ denotes exists. This is of course often true. However, it is just a special case of what people think. The underlying function of so-called *de re* expressions, like ‘Y’ in the formula (GR), is to express, not existence, but *speaker’s (intended) reference*. It is the person proffering (GR) who expresses *his* thinking of what *he* calls Y. The whole formula expresses that the speaker, *not* the person X, identifies what he calls Y with what X calls Z—regardless of whether Y exists or not. Obviously, to express this identification of his, the speaker himself also represents Z as Z. Terms in *de dicto* positions are points of referential *cumulation*.

To understand fully the nature of the acts of referring, we must understand (GR) in as general a way as thinking reaches. Doubtless, an episode of thinking is typically and primarily oriented toward the physical environment. Typically we think of real objects as having certain properties, and we succeed: the objects we think of exist and have the properties we predicate of them. Occasionally we make mistakes. Sometimes the objects we think do not have the properties we think them to have; sometimes the very objects we assume to exist do not exist. A comprehensive account of experience, thinking, and how language functions, requires acknowledgment that we need sometimes to refer to non-existents—as when we conceive failed plans to make certain things, or erroneously postulate certain entities to explain some events. An episode of thinking as a psychological reality is exactly the same whether what one thinks exists or not: the same brain events, the same representational structure. (Just imagine your thinking of a distant existent, which vanishes while you are thinking of it; yet its vanishing does not lame your thinking.) In sum, we must not equate externality of reference with existence.<sup>7</sup>

Thinking is impervious to existence. The sense of ‘refers to’ and ‘thinks of’ we need in (GR) lacks an existential commitment to what is thought of. We keep, on the other hand, an existential commitment to X,

the referrer. The sense of ‘refers’ we need is close to ‘purports to refer’. Yet this locution is valuable when an existential commitment is present. For convenience, we use ‘X refers to Y’, without implying that Y exists or suggesting that X thinks that Y exists. Thus, we subsume the problem of the reflexivity of self-consciousness under a more general problem of thought content.

Having access to another’s thought contents requires connecting those contents with the shared world. For this we must think ourselves those contents, whether we believe they are veridical or not. (This is referential cumulativity.) Thus, the general sense of (GR) can be pictured as follows:



where ‘S’ denotes the speaker’s identification of X’s Z with his Y, and each box represents a fragment of the world as conceived by the person mentioned immediately under it.

Statements of the form (GR) are little windows which the speaker opens to his hearer to let her peep into his (conceived) world; these windows reveal internal open windows on the said referrer X’s world. To speak (candidly) is thus to perform an act of revelation, of one’s world-window opening.

Internal, *de dicto* expressions are depictions of representations the speaker exhibits as SHARED by him and the person he speaks of. *De re* expressions just vent the speaker’s references. *De dicto* indirect speech has a Chinese-box structure, each box revealing the presented representations constituting thinking content—regardless of the existence of what is thought of.

A statement of form ‘X refers to Y as Z’ discloses to the hearer that the speaker is presented with thinker X as thinker X, with item Y as Y, with item Z as Z. Further, it reveals that the speaker takes Z to be an intersubjective thinkable. This is simply the cumulative character of the *de dicto* occurrence of Z in ‘X refers to Y as Z’. By equating Z with Y the speaker reveals, besides, that Z has an objective status in her world. This status need not be existence. Even if X in fact refers to the speaker’s Y as Z, Y may be only an intersubjective thought of item. In short, the exter-



nality of the moment *to Y as Z* in (GR) consists of its expressing the speaker's positing the *objective intersubjectivity* of Z as Y. She locates X's referred to item Z in her world or in the thinkable surroundings of her world. To be sure, the speaker may be in error in equating her Y with Z as thought of by X. It does not matter. The point is that she claims to have penetrated X's mind and have found Z there: Z is shared. Thus, the internality of the moment *as Z* of (GR) is the subjectivity of Z within X's experience, yet it manifests an accessible *intersubjective subjectivity* of Z. Again, this is the referentially cumulative function of *de dicto* construed terms in the Chinese-box structure.

On the assumption that the speaker makes true statements, what his world-windows show is part of the real world. Of course he may be in error; or he may be trying to deceive. The entity Y he locates in his world as being the same as Z as conceived by X may not be real. Clearly, what the speaker calls "Y," his Y, is—like Z in X's alleged act of referring—just a presentation typically (in the default case) pointing to the world. Thus, the item Y is no different in ontological status from the item Z. Both are aspects, faces, guises of thought of objects. In the typical episodes of thinking they are posited as real, that is, as existing aspects, faces, guises that compose unitary objects in the real world. In they are real then they have infinitely many properties not thought of in the thinking that the statement reports.

On the assumption that the speaker is telling the truth, we the hearers take it for granted that the aspect, face, guise Y put forward by the speaker for consideration is real. That is, we take it for granted that Y has a place in the causal order, and doxastically posit a whole infinite system of existing faces each *existing as being the same as Y*. Under this positing we can equate the speaker's Y with some item H in *our* world and say: "The speaker says of H, to which he refers as Y, that X referred to Y [it] as Z." Thus we become speakers, and assign to Y an intersubjective subjectivity: we and the (previous) speaker think of Y as such. This is the basis for our assigning to Y an objective intersubjectivity through our equating it with W. In its turn, Z gains a deeper intersubjective subjectivity: Z is then a shared item located in the innermost box of a three-membered Chinese box we have created.

Our H may not be real. We take it to be real, and our (conceived) world is the one we have beliefs about and act on. Metaphysically the basic fact is that we have NO access to an external point of view. All reference is from *our, one's* point of view. (As is well known, here lies the kernel of Kant's Copernican Revolution.)

The speaker's posited sameness between Y and Z need not be recognizable by the person X. It is not the strict identity of  $Y=Y$ ,  $Z=Z$ ,  $X=X$ .

As always, the elucidation of sameness is a serious matter.<sup>8</sup> Consequently, we must not prejudge the nature of the sameness involved in referring acts. We remain open to there being several different types of sameness required for the elucidation of storable mental representations, and for their occurrences yielding thinking presentations to a thinker.

Let us abbreviate ‘is the same as’ by ‘S=’. We allow that ‘S=’ may turn out to be a *sameness* variable. We reserve the sign ‘=’ for strict or genuine, necessary identity, self-identity, governed by a wholly unrestricted Leibnizian indiscernibility of identicals—including psychological and modal contexts. Hence, whereas some cases of sameness are bound to be contingent and empirical, strict identity is absolutely necessary. The symbols ‘X’, ‘Y’, and ‘Z’ have singular terms, not variables, as substituends. Clearly, ‘(Y=Z)’ does not imply (YS=Z). Let ‘YS=Z’ express the appropriate sameness relation suitable for each type of entity Y and the nature of what the speaker takes X to be thinking. Let ‘Z⟨X, t⟩’ denote the item Z as it appears in X’s world at time *t*, according of course to some speaker’s view of X’s world. Let ‘→’ express analytic or logical implication, and ‘↯’ its denial, these symbols having large scope. The basic logic of referring acts in general includes the following postulates, where *t* is a time:

- (Id.S)           YS=Z → (Y=Y & Z=Z)
- (GR.1)           At *t* X refers to Y as Z & XS=X’ & YS=Y’ → at *t* X’  
refers to Y’ as Z
- (GR.1.1)        At *t* X refers to Y as Z & X=X’ & Y=Y’ → at *t* X’  
refers to Y’ as Z
- (GR.2)           At *t* X refers to Y as Z & ZS=Z’ ↯ at *t* X refers to  
Y as Z’
- (GR.2.1)        At *t* X refers to Y as Z & Z=Z’ → at *t* X refers to  
Y as Z’
- (GR.3)           At *t* X refers to Y as Z ↯ YS=Z
- (GR.4)           At *t* X refers to Y as Z → YS=Z⟨Y,t⟩
- (GR.5)           At *t* X refers to Y as Z → at *t* X refers to Z as Z
- (GR.5.1)        At *t* X thinks that (. . . Z . . .) → at *t* X refers to  
Z as Z
- (GR.5.A)        At *t* X refers to Z as Z → at *t* X thinks something of the  
form (. . . Z . . .).

Here we leave it open whether some converse of (GR.3) holds. The contrast between (GR.1) and (GR.2)–(GR.5.A) is only one part of the contrast between the externality of the moment *referring to* and the internality of the moment *referring as*. The reflexivity of reference to oneself as Z, not necessarily as *oneself* (hence, not necessarily *self-reference*), is a special case of the externality of referring; it is, as already remarked, like the reflexivity of the self-eating tiger of the above example. (GR.4) connects internal thought content with an external thought of referent; (GR.3)

puts a limit to that connection. (GR.5) and (GR.5.A) are partial expressions of the presentational role of what is thought. They establish a necessary aspect of reflexivity in every act of reference, *a fortiori*, in the self-reference of self-consciousness.

Undoubtedly, there are three-position reflexive statements and facts of the form *X refers to X as X*. For instance:

- (2) The wisest man of Freiburg referred to the wisest man of Freiburg, to HIMself, as the wisest man of Freiburg.

By (GR.5) and (GR.5.A), it suffices that the wisest man of Freiburg thinks, for example, this: The wisest man of Freiburg will exert great influence in the cultural changes in the Albert-Ludwigs-Universität. Palpably, in his mere thinking episode reported in (2) the wisest man of Freiburg does not refer to HIMself as *himself*. This is so even if he knows that he himself is the wisest man of Freiburg. This piece of knowledge may indeed be operative and underwrite his thinking what (2) reports in a context of impersonality. Obviously, then, three-position reflexivity of referring is *not* necessary for *self-consciousness*; it is actually inimical to *self-consciousness*.

Now we come to the third prong, 'as (*qua*) *oneself*,' of our initial formula of *self-consciousness*:

- (I\*) In an episode of *self-consciousness*  
ONE refers to (thinks of) ONEself *qua oneself*.

We turn to the second meaning of the suffix '-self', that is, the meaning of the italicized pronoun '*oneself*'. In (I\*) this is a relative pronoun having 'ONE', *not* 'ONEself' as its antecedent. Logically, this composition is iconic: '*one-*' + '*-self*'. To the extent that '*oneself*' has an antecedent it has the character of this antecedent. In our schematic formula (I\*) the antecedent 'ONE' is a variable; hence, '*oneself*' has a variable aspect: the prefix '*one-*' is the image of the variable 'ONE'. The substituends of '*oneself*' have as antecedents the corresponding substituends of 'ONE'. That is, the substituends of '*oneself*' co-refer with, because they refer back to, the substituends of 'ONE'. The substituends of '*oneself*' in (I\*) are reflexive pronouns used to attribute first-person reference in appropriate indirect speech constructions: '(he) himself', '(she) herself', 'themselves', 'yourself', 'myself'. For example:

- (3) The author of "Self-reference and Self-awareness" is thinking of HIMself that *he himself* wrote *Self-Knowledge and Self-Identity*.

Here the prefix '*he him-*' refers back to its antecedent 'The author of "Self-reference and Self-awareness"'.

All the mentioned expressions in the uses under consideration share a semantic component, which is carried by the suffix ‘-self’. Hence, the locution ‘*oneself*’ (like its substituends) in the contexts ‘ONE refers to ONEself as *oneself*’ depicts, is a proxy for, a first-person reference attributed to *one* (or the entity denoted by its substituends). Here ‘*oneself*’ is what I have called a *quasi-indicator*, more precisely, a quasi-indexical variable.<sup>9</sup>

Consequently, first-person quasi-indicators have a crucial internal, *de dicto* semantico-pragmatic strand: they depict the subjective internal self-reference peculiar to self-consciousness. They also have a most important external, *de re* syntactico-semantic dimension: through their connection to their antecedents they express an intersubjective objectivity of first-person reference. (Here is another reason why I do not like the standard *de dicto/de re* contrast: it is supposed to be exclusive and exhaustive; thus conceived it makes it impossible to understand the function of quasi-indicators.)

For perspicuity let’s append a star, ‘\*’, to an expression that functions quasi-indexically: as a depiction of an indexical reference. The star in ‘one\*’ or ‘oneself\*’ (‘he\*’ or ‘himself\*’) thus replaces the italics in ‘as *oneself* (*himself*)’. Recall that in ‘ONEself’ the suffix ‘-self’ simply signals that ‘ONEself’ and ‘ONE’ are the self-same variable. Thus, our formula (I\*) has the sense of:

(I\*.1) In an episode of *self-consciousness* X refers to X as himself\*.

As noted above, the quasi-indicator ‘himself\*’ is composed of ‘him-’, which refers back to ‘X’, and the self-presentational content ‘-self\*’. To signal this representational duality of quasi-indicators, we could put our formula of self-consciousness canonically as follows:

(I\*.F) In an episode of self-consciousness  
a thinker X refers to X as [X]\*.

The bracketed component ‘[X]’ represents the external anchoring of the quasi-indicator and of the depicted indexical reference; the brackets signal that this anchoring is external, a part of the speaker’s location of X’s *I* in his own world. The star ‘\*’ signals the internal depiction of a certain indexical reference. As remarked, the pseudo-internal ‘X’ in ‘[X]’ has as antecedent the first occurrence of ‘X’, which denotes the referrer. Hence:

(I\*R.1) The formula “Y refers to X as [X]\*” is nonsensical, and so is the formula “X refers to [X]\* as Y.”

(I\*R.2) The formula “X refers to X as [X]\*” is an instantiation of the formula “X refers to Y as [X]\*.”

Obviously, (I\*R.1) subsumes the impossibility of referring to another in the first-person. On the other hand, (I\*R.2) records that one can attribute

first-person references to somebody else whom one thinks to be the same as oneself\*. The attribution of a first-person reference is a third-person quasi-indexical reference. Doubtless, rehearsing one's mis-self\*-identification with another does require first-person reference to oneself *not* to that other.

Recall that we are using the idealized term 'refers' with an existential commitment to the referrer. Hence:

(I\*R.3) At  $t$  X refers to Y as  $[X]^*$   $\rightarrow$  X exists at  $t$ .

(I\*R.4) At  $t$  X refers to Y as  $[X]^*$   $\rightarrow$  at  $t$  X assumes (knows) that  $[X]^*$  exists at  $[t]^*$ .

NOTE. The quasi-indicator 'at  $[t]^*$ ' is expressed in English as 'then'.

(I\*R.5) At  $t$  X refers to Y as  $[X]^*$  &  $XS=A$  &  $YS=B$   $\rightarrow$  at  $t$  A refers to B as  $[A]^*$ .

(I\*R.6) At  $t$  X refers to Y as  $X \rightarrow$  at  $t$  X refers to X as  $[X]^*$ .

(I\*R.7) At  $t$  X refers to Y as  $[X]^*$   $\rightarrow$  at  $t$  X refers to Y as X.

Principle (I\*R.6) formulates in part the point, illustrated with example (2), that three-position reflexivity is inimical to self-consciousness. Both (I\*R.6) and (I\*R.7) are illustrated by (3) above and (4) below:

- (4) The author of "Self-reference and Self-awareness" is thinking of the author of "Self-reference and Self-awareness" that the author of "Self-reference and Self-awareness" wrote *Self-Knowledge and Self-Identity*.

Clearly, (3) neither implies, nor is implied by, (4). This failure of implication holds not only for episodes of thinking, but also for dispositional states, e.g., believing. This is confirmed by a recent report that our author is in doubt about the existence of the book mentioned above, and has forgotten his ever having written an essay with the above title. (How the author of such great works can forget his writing them is not a conceptual matter, but an entirely different, biographical, question.)

In general, for any equation ' $j$  S = the author of "Self-reference and Self-awareness,"' there is a pair of sentences (3. $j$ )–(4. $j$ ) paralleling the pair (3)–(4). The author of the essay can think (even believe) that  $j$  wrote the book without thinking (or believing) that he\* wrote it.

First-person reference is not just internal to a thinker's thought contents. It is unique even within the internal contents of a thinker's thoughts. Each one makes her own demonstrative (indexical) references to items in her experience. Yet one can think demonstratively of an entity X, which one encounters in a certain experience (for instance, by looking into a mirror), without thinking that X is *oneself*. Furthermore, a thinking entity, still without identifying what it sees with itself\*, can address the entity it sees in the mirror, expressing thus some second-person thought: "You look terrible. You should not present yourself in public like that. . . ." Therefore, "That (person) [she, he] is F" and "You are F" are entirely different

thought contents from “I am F.” Thus, the internality of third-person demonstrative reference and also of second-person reference may have a *de facto* external reflexivity.

First-person reference is necessarily reflexive. This feature is signaled by our formula ‘ONE refers to ONEself as *oneself*’. The expression ‘*one-*’ in ‘*oneself*’ (and, canonically, ‘X’ in ‘[X]\*’) does co-refer with ‘ONE’ (that is, the first occurrence of ‘X’). This reflexivity forces first-person reference to harpoon an external reality. This is partially captured by (I\*R.3) and (I\*R.4). By means of (GR.5.1) we have:

(I\*R.8) At  $t$  X thinks (. . . I . . .)  $\rightarrow$  There exists an entity  $u$  such that  $u S = I\langle X, t \rangle$ .

This principle is valid for every speaker. Therefore, it secures a referent of the first-person pronoun in the shared world, the real world. The speaker who uses ‘I’ is no exception. Hence, for each thinker the following hold:

(I\*R.8.1) At  $t$  I think (. . . I . . .)  $\rightarrow$  There exists an entity  $u$  such that  $u S = I\langle I, t \rangle$ .

(Id.I\*\*)  $u S = I\langle I, \text{now} \rangle \rightarrow u S = I$ .

(These postulates constitute the core of Descartes’s *cogito*.)

Let us turn now to the particular *internal* content one thinks in self-reference. Doubtless, the external target of self-reference is ONEself. One always succeeds in harpooning it. But what is the presentational representation with which one harpoons ONEself? Evidently, it is what a use of the first-person pronoun presents to its user: an *I*. This is a unique and irreducible representation. Its irreducibility is precisely what (I\*R.6) and (I\*R.7) formulate, what the non-implication between (3) and (4) shows. The first-person pronoun has a general meaning that determines a framework within which, on each occasion of its use, a speaker confronts a unique and personal representation of the reality he or she is. Every one thinks of himself as an *I*. The others can represent that personal representation, not just the private presentational role it plays in the given episode of thinking, by means of a quasi-indicator. Consider:

(5) X thinks (outloud for our convenience): I love Elly.

A hearer asserts:

(6) X thinks that he\* loves Elly.

By making statement (6) the speaker opens a window to her world. Again, this window reveals a window open to X’s world. This time, however, we do not see an item Z in that part of X’s world shared with the speaker. We see a faithful painting—*himself*—of a literally unshared item in X’s world—what X calls “I.” Clearly, ‘himself’ is internal to the con-

struction (6), somehow presenting to the hearer the presentational representation X confronts. Nevertheless, it does so indirectly. Evidently, X does not use the third-person pronoun 'he himself' to refer to himself as himself: he uses the first-person pronoun 'I'. Only his uses of 'I' refer to him—as himself\*, of course.

In brief, the indexical uses of the first-person pronoun have thus three central features: (i) they have an ineradicable privacy of presentational representation; (ii) they infallibly harpoon an existent beyond the representation, and (iii) they intrude as also infallible the very existence of the presenting representation, that is, the thought of *I* itself. (The other types of indexical have properties similar to (i)-(iii). The only exception is demonstrative reference to experienced items, which lacks property (ii). Statements of the form "This is F" secure an internal referent, the subject *this*, but do not always harpoon an external referent.)

The uses of the quasi-indicator depicting first-person reference have therefore five crucial properties: (i) they occur in internal, *de dicto* positions as depictions of *I*-reference; (ii) they are literally third-person representations of such references; (iii) they involve an internal, reflexive existential commitment: they present the presentational depictions as existing; (iv) they involve a trans-internal existential commitment to their depicted *I*'s; (v) they carry an external existential commitment to the referents of their antecedents. *I*-presentations, like all other indexical representations, exist only in their presentational function. On the other hand, our quasi-indexical representations of others' *I* representations are interpersonal and lasting, and can have repeated presentations in diverse thinking episodes. This is the constant content signaled by our quasi-indexical star in 'himself\*' and in '[X]\*'. Consequently, indexical statements or propositions are different from the quasi-indexical propositions, or thought contents, that correspond to and depict them.<sup>10</sup>

But what exactly is that *I* content one is presented to in episodes of self-consciousness? That it is so peculiar and unique is further brought home by the fact that there is no special characteristic that one has to think that one possesses in order to think of oneself as *I*. Certainly, one *qua I\** does not classify oneself as a self, a person, or a thinker—let alone as a human being, female, or whatever is true of all entities capable of *self*-consciousness. To illustrate, a small child at about the age of two can make perfect first-person references fully lacking knowledge involving those categories. In general:

- (I\*R.9) Where *being F* is not simply *being myself*, or *being thinking*, or *existing*, or their joint logical implicates, all propositions of the form *I am F* are synthetic *a posteriori* in Kant's sense, i.e., non-redundant (non-tautological).

There is just no criterion one can apply to determine whether one is an *I* or not. One simply is an *I*. This primitive fact is primitively and immediately apprehended by a thinker who is an *I*.

Let us look at the situation from another direction. As noted at the beginning, the dominant standard approach in the study of language and mentality takes dialogues as basic data; this has led to the adoption of an external, hearer's perspective. Yet certainly in the case of self-reference we must adopt the position of the speaker and her thinking references.

The following is certainly a correct rule governing the use of the first-person pronoun:

- (K-I) In any statement in which an indexical, singular use of a token of the first-person pronoun occurs, that use (or that token) designates the speaker of the statement.

This is a marvelous rule for *communicating* self-consciousness, not for *experiencing* self-consciousness. We can ignore the fact that one often thinks without making statements. We may suppose that at some level in the language of the brain we have personal pronouns; we may understand 'speaking' broadly enough to mean the production of symbols, whether overtly in a public language or covertly in the language of the brain, which symbols embody thought content. On these assumptions, (as Plato said) speaking is thinking, and thinking is speaking. By these semantic stipulations we secure that a hearer who understands a used sentence containing the first-person pronoun understands, concerning this pronoun, no less than what (K-I) requires.

Nonetheless, a hearer has to understand *more* than what rule (K-I) prescribes. The hearer has to understand the internality of the references the first-person pronoun has been invented to express. Rule (K-I), if not in the pedagogical process, at least in the ontological order, rests on the following rule:

- (I\*.I) A thinker who makes a statement with a token of a sentence containing one token T, or more, of the first-person pronoun, indexically *used*, refers with T to HIMself *as himself*.

Furthermore, if all thinking is speaking, a *speaker* cannot act from (the conception of) rule (K-I) or (I\*.I). He will have to conceive *himself\** as the author of a first-person statement not yet made, or made in the very act of thinking something in the first-person way. But then how can he identify *himself\** with the maker of that statement? This identification would simply be embodied in a use of the first-person pronoun that precedes the application of the rule.

Moreover, language must be spoken, indeed, before it is understood by a hearer. If thinking is speaking to oneself, then to hear oneself one must sometimes speak without hearing oneself. The fact is that for a



speaker his\* referring to the maker of a statement with the first-person pronoun cannot be a criterion for his use of the first-person pronoun.<sup>11</sup>

Still what is the content one thinks as *I*? This is a serious question that requires a separate study. In part IV is the briefest outline of the answer.

We must fasten to the fact that each *I* is a primitive and ephemeral internal representation of the thinker as the subject of occurrent internally, presentationally convergent experiences. Because of (I\*R.9) this representation contains no description or attribution of a property. It is merely the brute unanalyzable presentation of the internal unity of those experiences with their successful harpooning of something in the real world otherwise unknown—whatever it may be. Hence an *I* content is in general the presentation of an operational unification of a set of structures that unify the particular contents of experience. To think an *I* content is to execute a unification of presented representations and to be presented with that execution.

### III

#### THE WORLDLINESS/NONWORLDLINESS

#### TENSION OF SELF-CONSCIOUSNESS

I. THE UNWORLDLINESS OF THE I'S     The internality of what we think and refer to in indexical first-person reference is fundamental and absolute. Overwhelming; tantalizing. The content of the reflexivity of *self*-consciousness is wholly internal and peculiar to the thinker's thought contents. That is: (I\*\*R.1) An *I* exists subjectively only, within the contents of episodes of *self*-consciousness.

This conforms to Descartes's immediate hesitancy after confronting his Evil Demon (or Wilfrid Sellars's Mad Scientist).<sup>12</sup> He claims cautiously (where the capitals express my added emphasis) that:

*De sorte qu'après y voir bien pensé, et voir soigneusement examiné toutes choses, enfin il faut conclure, et tenir pour constant que cette proposition: je suis, j'existe, est nécessairement vrai, TOUTES LES FOIS QUE JE la prononce, ou que JE la conçois en MON esprit.*<sup>13</sup>

Yet there is an exegetical problem with this *dictum* of Descartes's. Cautiously, he acknowledges that *I exist* is true, indeed, necessarily true, WHENEVER the proposition is thought by him. This leaves open that it may be false when it is not thought. However, then there is no *I* to think that proposition. This would intimate that perhaps each time *I exist* is thought a fresh new *I* is thereby brought into existence. All of this is in accordance with the internality of the *I*'s. On the other hand, Descartes's *I exist* is true every time I (Descartes) (a translation of 'JE') think so. What

about this I that thinks at different times? Descartes of course did not bother with this. He simply took the capitalized I to be what his argument establishes. The immediately following paragraph begins: “*Mais je ne connais pas encore assez clairement ce que je suis, moi qui suis certain que je suis.*” These expressions ‘*je*’ and ‘*moi*’ simply denote the JE of the preceding citation.

Descartes seems to be urging principles that in a more general form can be put as follows:

- (I\*\*R.2) For every I-proposition: (. . . I . . .) → I (the thinker) exist.
- (I\*\*R.3) For every I-proposition: (. . . I . . .) → I am thinking or I have thought.

I submit that the necessity of the truth of *I exist* that Descartes proclaims is just the necessity of the above implications. These are necessary truths in the logic of *self-consciousness*, or, better yet, in the logic of the *I*'s. These Cartesian principles are, however, compatible with a non-Cartesian, but quasi-Humean view of the *I*'s. This is the view mentioned above: the *I*'s are merely subjective entities, whose existence is exhausted in episodes of *self-consciousness*. On this view the essence of an *I* is, as for Descartes, thinking (*cogitare*), but, anti-Cartesianly, an *I* is not a *res cogitans*.

2. THE REQUISITE WORLDLINESS OF THE I'S      The more internal the *I*'s are the more non-real they become. Now, the pinnacle of human living is to live an examined and carefully planned life: a life that consists of the unfolding of a system of intentions. These are however first-person thought contents.<sup>14</sup> Consider a somewhat revised example I have used before:<sup>15</sup>

*The Quine Award to be Given to Sellars in the Chisholm Auditorium*

- (7) The Plantinga-Grünbaum Society for the Philosophy of Mind has unanimously approved that its President present Wilfrid Sellars with a special band containing an appropriate gilt inscription, an elegant diploma, and a check for US \$1,000,000, the three together embodying the Willard V. O. Quine Award for Philosophical Accomplishment. The ceremony is to take place at the Roderick Chisholm Auditorium of Brown University on May 12, 1989.

In the situation under consideration in (7) for somebody to be in a position to attempt to act intentionally, it is necessary that he/she places the entities in question in his/her experience at the proper time. The agent must answer questions about the relevant ceremonial identifications. If an agent is to carry out the ceremony voluntarily and intentionally, that agent must endorse the following answers to such indexical questions:

- (8) *Indexical identifications*
  - (a) *I* am the President of the Plantinga-Grünbaum Society.
  - (b) *This here* is [you are] Wilfrid Sellars.

- (c) *This* is the Chisholm Auditorium of Brown University.
- (d) *That* is the P-G Society band of the Willard Quine Award.
- (e) *That* is the P-G Society diploma of the Willard Quine Award.
- (f) *Today* is May 12, 1989.
- (g) *Now* is the time of the Award conferral ceremony.
- (h) *Here* is the location where the ceremony is scheduled to take place.

With these propositions in his/her doxastic repertoire, a person can become the agent to bestow The Quine Award on Sellars. Indexical equations (8)(a)–(h) constitute the bridge on which the agent-to-be can move from endorsing the external singular proposition (7) to thinking endorsingly its corresponding *indexical deontic proposition*, to wit:

(7\*) *I* ought to give to *this* man *this* band and *this* diploma *here* now.

The success of our intentions and our complying with our duties requires that we act in one and the same world as we cognize. Intentional and voluntary action hinges on there being indexical internal contents that the agent takes to be true of the world. This is what (7\*) illustrates. The connection between (7) and (7\*) needs the mixed indexical-nonindexical equations (8) that link our experiences to the world at large. The sameness proclaimed by the equations (8) is a transcategorical sameness between the indexical items present in experience and their harpooned targets in the world.

In brief, such absolutely indispensable mixed equations like those in (8) have two obverse functions. From the side of subjectivity, they provide the indexical particulars—the *I*'s, the *this*'s and *that*'s, the *now*'s and *here*'s—with worldly anchors. They all have, therefore, their deep-grounded worldliness. From the side of objectivity, the mixed equations provide the external entities that are experienced—the agents, the presidents of societies, the buildings, the Sellarses, the Quines, the Chisholms, the times, and the places—with their appropriate distinct experiential handles. Obviously, the mixed character of those equations presupposes that the equated subjective and objective items are different individuals: they certainly differ in ontological status. What sort of sameness is that?

3. THE SOCIAL WORLDLINESS OF THE I. The necessarily worldliness of the *I*'s is valid for an agent, even if he is a lone and solipsistic agent. Of course, example (7) is not solipsistic. Nevertheless, (7) considers the situation (almost) exclusively from the point of view of the one who presents the award to Sellars. The participants are also sentient and thinking beings. For them to cooperate with the agent they must ascribe intentions and beliefs to the agent and to one another. Hence they must place in the shared world the others' *I*'s (and *this*'s and *that*'s, *now*'s and *here*'s). Then for the

benefit of others, which revolves into one's benefit, the logic of *self*-consciousness requires this postulate of *I*-embodiment:

- (I\*\*R.4) No *I* is an ontological island unto itself. There must be infinitely many unique intersubjective characterizations F-ness such that:  
One is (the same as) the person F,  
e.g., I am (the same as) the person F.

4. MAJOR PROBLEM: ONE TYPE OF SOLUTION      One serious problem for any theory of *self*-consciousness is to provide an account of *I*-hood that reconciles, assuages, or dissolves the deep-rooted tension between the non-worldliness and the worldliness of the *I*'s. Non-worldliness arises from the internality of *self*\*-reference; worldliness springs forth from the externality of SELFreference, which necessitates the embodiment of each *I*.

It seems that the most economical solution to the worldly/nonworldly tension about *I*'s requires a theory of sameness that recognizes different sameness relations. One simplification is this: The way in which every *I* must be the same as a physical object others can identify is the same way in which the thinker HIMself is the same as each of his ephemeral *I*'s. Clearly, this sameness is contingent and MUST be carefully distinguished from necessary samenesses, especially strict identity. Guise Theory is an already available theory of samenesses.<sup>16</sup>

## IV

### THE HIERARCHY OF REFLEXIVITY OF SELF-CONSCIOUSNESS

#### I. THE NETWORK OF STRUCTURAL NEGATIVITIES COMPOSING I-SCHEMATA

Once again, an *I* is what one thinks as an individual with an indexically used token of the first-person pronoun. The categorization "as an individual" is meant to capture the common core across the different inflections of the pronoun. These inflections represent important distinctions. For instance, a token of 'me' used as a dative conveys information crucially different from that conveyed by an accusative token of 'me'. Both differ drastically from the information borne by a token of the genitive 'mine', and even more from that riding on a token of the possessive adjective 'my'. These are local semantic contrasts internal to our concept of the first-person. They represent subtle structures enveloping each *I*, within which a given *I* enters into particular relationships.

We face four related ontological questions:

First: Does each token of the first-person pronoun denote a unique *I*?

Second: If not, do all the tokens of the first-person pronoun issuing

from one physical source denote one and the same *I* throughout the history of that source?

Third: If so, can the tokens of '*I*' emanating from different sources designate one and the same unifying *I*?

Fourth: If so, do *all* the tokens of the first-person pronoun designate one omni-comprehensive *I* that underlies the different utterances, regardless of the uttering source?

The answer to the first question must be a categorical and emphatic "No!" The tokens of the first-person pronoun in one piece of discourse expressing a unified train of thought, or experience, at least *normally* denote the same *I*. This *I* unifies the experience whose contents are deployed in that piece of discourse. Nevertheless, the unity of the experienced content is a logically necessary condition of, but not a sufficient condition for, the uniqueness and the unity of the experiencing *I*. In principle, the tight unification of a rich experienced content could be experienced by different *I*'s. We could have in such a case *psychological Siamese twins*. Given a physical or physiological basis for episodes of consciousness such twins do not seem entirely unfeasible. Two brains may for certain experiences be connected to one common source of the experience. Can, for example, two Siamese twins sharing one leg and one arm feel exactly the same pain on the shared leg? Or can two different personalities embodied in the same brain, or in different parts (for example, different hemispheres) of the same brain, share the same numerical experience? Of course here we move within an area where stipulations suitable for economic description of our stable world are available. Yet different meaning postulates may provide a better terminology for adequate theorizing. The problem is more for psychiatrists. They treat patients who together present an extraordinary diversity of experiences. Their problem is to make enough distinctions that could enable them to deal with all their cases. Short but thick terminologies that express narrow theories are often torn apart by scientific progress and technological developments.

The answer to the second question is generally affirmative.<sup>17</sup> I am one of those who prefer a negative answer. We shall consider this below.

Those who answer "yes" to the second question should in general be flexible enough to answer "yes" to the third question. In the less committal cases the answer to the third question depends on the adopted criteria of personal identity. To the extent that a person can alter its (her, his) physical embodiment, given a certain diachronic unity of experiences, the same lasting *I* might be said to underwrite that unity. On the other hand, an affirmative answer to the third question is a basis for an affirmative answer to the fourth question.

Let's return to the *I*-structures that mold the contents of *self-consciousness*. We seek further *I*-enveloping structures beyond the local ones carried

by pronominal declension. They can be gleaned by charting the types of properties and relations characteristic of *I*'s. Each such type is a domain of possibilities open in principle to all *I*'s. Here we will not carry out this cartographical investigation. Our present concern is confined to the internal *I*-structures, bundles of which constitute *I*-schemata. These structures have certain independences; hence, they need not be present all at once. There is, thus, a hierarchy of *I*-schemata. Consequently, the *I*'s can be partially ordered in accordance with the ordering of the *I*-schemata constituting them. Thus, the *I*-schemata provide blueprints for egological development. Under a conception of normality, the *I*-schemata furnish also a master chart against which to check a person for possible *I*-disturbances.

The internal *I*-structures are *I*-strands hinging on contrasts between what one is *qua* oneself\* and something one is not. Alternatively phrased, an *I*-strand is the polar negation of something intrinsically non-*I*. Because of the polarity, the negation is dichotomous.<sup>18</sup> Thus, an *I*-schema is a complex of *negativities*. To gain a concept of *I*, is to acquire the capacity to pick up immediately *instances* of one or more of those polar negations. This instantiation is of course picked up by thinking of ONEself *qua* oneself\*. Here is a fundamental ingredient in the peculiar reflexivity of *self*-consciousness.

To think of ONEself as oneself\* is to think of something *intimated* to be, *felt as it were*, the opposite of each of such and such N's—for the appropriate values of 'N'. I have chosen the words 'intimated' and 'felt' advisedly. We have already seen how to think contents of the form (. . . I . . .) is to have a primitive apprehension of the subject one calls "I," not mediated by any identification procedure. That basic truth remains unaltered. Hence, to apprehend oneself\* in the content (. . . I . . .) is to apprehend a manifold of polar negations of the different non-*I*'s as one fully unified manifold of *I*-strands. Furthermore, it is to apprehend that manifoldness in a non-conceptual way, as a sensory content. Conceptually one apprehends only the unity. The apprehension of the manifoldness underlying the unitary *I* consists partly of feeling that sensory-like content. Intellectually, such an act of apprehension sets in readiness a hierarchical manifold of propensities to think, especially to think believingly, appropriate ranges of propositions. These ranges are demarcated by the felt negativities. The elements of those ranges are determined by the personality of the thinker, the context of the thinking episode, including her surface purposes, and her preceding trains of thought.

*Self*-consciousness is, thus, erected on a reflexive sensory consciousness. This reflexivity is not reflective (in the sense of reflecting on something); it is of the ONEself type of reflexivity discussed in part II above. We will see that *self*-consciousness is mounted on other forms of non-reflective consciousness. Also *self*-consciousness occurs on a doxastic pedestal com-

posed of propensities to think or to rehearse belief. This pedestal is a hierarchical arrangement of those propensities along an ordering of degrees of triggered readiness for activation. The uppermost tiers are propensities to think, which may be also activated, appearing through propositions also present to consciousness; some propensities will manifest themselves in the penumbra of consciousness; others lurk behind perhaps as merely felt in sensory consciousness. Most of the tiers of the pedestal are utterly unconscious. The bottommost ones are metaphysical; yet they are not so much beliefs that can on occasion be brought to consciousness; they are, rather, deep-seated takens-for-granted, built in the hardware of the inferential or computational mechanisms that underlie the embodying of thinking episodes. These cannot (physically, or psychologically) be brought to consciousness. Some of them may actually be unformulable.

To elucidate the preceding paragraphs we must discuss, even if briefly, the nature of doxastic pedestals. Before we do that we must, however, say something about the network of the negativities composing the *I*-strands.

#### Chief *I*-strands

1. The contrast *I-this/that*
2. The contrast *I-they* [the external objects of the world]
3. The contrast *I-he/she* [this/that (thinker)]
4. The contrast *I-they* [the others]
5. The contrast *I* [believer/knower]-*I* [agent]
6. The contrast *I-he/she* [this/that (person: sentient-thinking-and-acting individual)]
7. The contrast *I-you*
8. The contrast *I-we* [partners at a conversation]
9. The contrast *I-they/we* [the members of one's community]

Remark 1: Here I have been inspired by Ferdinand de Saussure in conceiving a meaning of an expression as a contrast between a usage of the expression and the usages of other related expressions within a family.<sup>19</sup> Each contrast of usage is a semantic strand. To me Saussure's general thesis seems to be correct, but it is particularly suitable for the semantics of the first-person pronoun. Of course here the usages and meanings of words interest us only as avenues for approaching thought contents. The above nine contrasts are contrasts in thought content, in what we find in experience, and what we believe to be in the world.

Remark 2: I have adopted the hyphen in between the contrasted expressions from Buber. He contrasted "the two meanings" of 'I': *I-It* and *I-Thou*.<sup>20</sup>

Remark 3: The different *I*-strands listed above divide the full-fledged concept of *I* along different axes. The endpoints of these axes are: (i) what is being experienced vs. what perhaps (although not experienced) belongs nevertheless to the world at large; (ii) what is internal to the mind vs. what

is external; (iii) being vs. not being a person or a mentally endowed individual; (iv) cognizing the world vs. acting on it, which rests on the axis believing vs. intending; (v) being a thinking-acting individual vs. being a member of a community. Undoubtedly other axes must be added for a full account of the *I*-schemata and their strands.

Remark 4: The nine *I*-strands enumerated above intermingle in many different ways. This yields a spiralic process of steady enrichment of one's concept of *I*. The mingling principles are of different sorts. Some are required by the kind of world we find ourselves in, with its physical, chemical, and biological natures, but also with its social organizations and other cultural products. Some principles of mingling lie deeply seated in the thinker's ontologico-psychological makeup, which manifests itself in the thinker-agent's metaphysical postures. For example, animists mingle the *I*-strands in such a way that everything has its own internal *I*. Solipsists find the total world at large to be merely experiential content of the only accessible *I*. Deists promote the whole of reality to the veridical content of an all-embracing *I*. Pantheists are animist deists. Mystics do all sorts of things; some raise themselves up to the status of partial *I*'s within an all-encompassing one.

Remark 5: The order in which the *I*-strands are listed is logical from the perspective of the experience of thinking. It is not meant to prejudge the philosophical disputes about the possibility, or viability, of solipsistic consciousness. It is not intended to cast any aspersion on, or endorse the converse of, the semantic socialism now fashionable.<sup>21</sup> It is not meant to decide in advance the outcome of controversies (philosophical, pedagogical, or otherwise) about the learning of language or the acquisition of concepts.

Remark 6: The network of negativities is a network of structures. They have to be filled in with special contents. These are empirical and metaphysical beliefs, or even deep-seated takens-for-granted.

2. THE HIERARCHY OF MODALITIES OF CONSCIOUSNESS WITHIN SELF-CONSCIOUSNESS As noted the existence of *self*-consciousness rests on a complex doxastic pedestal. This in its turn pivots causally, so to speak, on an iceberg of unconscious processes. The contents of *self*-consciousness are not a uniform monolith somehow above the line separating what is in the light of consciousness and what is submerged in the dark water of the unconscious. Those contents have, rather, a hierarchical structure. This is not just a hierarchical order of those contents; it is also a hierarchy of modalities of consciousness. *Self*-consciousness subsumes all other modalities of consciousness. This is another dimension of the characteristic and peculiar reflexivity of *self*-consciousness. In general, consciousness



is cumulative and suffusive. *Self*-consciousness is just more so. The unitary structure of consciousness proceeds upwards, but its illumination flows downwards. The higher, encompassing degrees alter the nature of the lower degrees.

The internal hierarchy of consciousness within *self*-consciousness includes the following levels:

1. sensory content, conceptually inarticulated
  - (a) bodily
  - (b) worldly
2. I-less articulated content pertaining to
  - (a) external objects
  - (b) bodily content
  - (c) occurring mental acts
3. I-less focal consciousness, the core of which is a complex of perceptual judgments
4. I-owned content articulating the contrast between Self and Object
5. I-owned content articulating intentional agency
6. I-owned content articulating the contrast between Self and others
7. I-owned content articulating an interaction between Self and *you* as well as between self and absent persons.

The maximal degree of *self*-consciousness has an *I* as its focus and anchor. This fills in the *I*-schema in which all the relevant negativities intersect and have their proper empirical and metaphysical content.

The integration of a manifold of contents as the contents of a unitary *self*-consciousness proceeds from the lower levels up. That is, we cannot explain the unity of an episode of consciousness in terms of an *I* to whom that episode belongs. This can be established empirically: there are episodes of *I*-less consciousness. It is not just that there is a quiet *I* in the background, which can at any moment come into the open.<sup>22</sup> At this point, the positing of an enduring *I* as the underwriter of the experiences of one and the same person at different times becomes handy. (Recall the *second* question above in section IV.1.) This positing must, however, be grounded on some evidence or reasons. (See note 5.) No matter. The *I*-less consciousness of some lower animals obtains without being owned by any *I* that can come to the light of *self*-consciousness.

All consciousness is cumulative and suffusive. Accordingly, I submit that in gaining full consciousness after a deep sleep we go through all the levels of consciousness as different stages of the process of waking up. As I see things, since each *I* exists only during and within an episode of *self*-consciousness, waking up is really a case of ontogeny recapitulating phylogeny. Anti-recapitulation occurs in cases of falling asleep slowly.

*Level I* is a theoretical posit. We assume that there is a manifold within sensory consciousness which includes no awareness of objects or events. The primary consciousness is perceptual. We may assume that the

sensory manifold is, underneath consciousness, partitioned into masses belonging to the different types of perception.

At *level 2* there is conceptualization of some sort. The minimal sort is what I have called *zero-consciousness*.<sup>23</sup> This is the confrontation of a perceptual field that contains as distinct points some perceptual states of affairs (or propositions) as unanalyzed units, with no logical structure, not even a subject-predicate one. I theorize that this is the kind of consciousness exemplified by animals that respond to differences in color, shape, distance, without any consciousness of objects as such. However, I know so little about animals that I cannot offer a well studied example. In higher degrees of *level 2* some of the presented states of affairs (propositions) receive some structural analysis. The subject-predicate structure enters the scene. This consciousness is *I*-less and not focal.

A very interesting case of *level-2* consciousness is so-called blind-sight. Some persons who have lost part of their cerebral cortex connected to vision illustrate what to many specialists seems to be a paradoxical situation. While declaring that they do not see, such persons are nonetheless able to respond correctly to questions that seem to require clear visual perception. A certain patient—named D. B. in the literature—is particularly interesting. He had part of his brain removed but had lost only his left visual field. In the experiments, items were placed in his visual field. He insisted that he\* (himself) did not see anything in his left field. He engaged in “guessing,” as he called it. He had a perfect score in thirty “guesses.” D. B. sometimes reported having a “feeling” that certain figures he said he\* did not see had certain properties, which they in fact had.<sup>24</sup> Similar studies have been carried out with monkeys. Apparently monkeys and men with blind-sight can improve the strength of those “feelings” and also their powers of discrimination of “unseen” objects.

Some philosophers prefer to interpret the cases of blind-sight as cases of no perception at all. To them they are evidence that consciousness is a well-entrenched fiction, which science will eventually vanish.<sup>25</sup> Others see blind-sight as cases of vision, that is, visual consciousness without *self*-consciousness.<sup>26</sup> This is the view I have adopted.

As I see it, in blind-sight the integration of the visual contents is, first, incomplete. Then only the fully integrated part of those contents acquires an owning *I*. The contents of blind-sight remain at *level 2*. Now, the patient learns to link the *I*-less contents of his left field to those of his *I*-owned contents of his right field. This produces a partial lopsided integration of his *level-2* contents. There is now an *I* that owns the new structured experience. This, however, remains mixed because of the partial integration. The contents in D. B.’s left field become accessible to him as if they were seen, but they continue being as if they were not seen *by him*.

Consciousness is cumulative. Hence, *level 2* is present in the higher levels. Most of what we see surrounds the part of the visual field where we focus, and we see it within consciousness of *level 2*. Sometimes we store such contents in memory in the form of mechanisms for producing visual images. These mechanisms can be turned on by using the memory of the focal contents of the original field, so to say, as handles or push-buttons. Then the produced image can be inspected by changing the focus. This way one can *ex post facto* promote contents of consciousness from *level 2* to *level 3*. This posthumous promotion is feasible because of the integration of the contents of these two levels within a consciousness of *level 4*. This provides the stage. The promotion itself, executed by a voluntary action, occurs within consciousness of *level 5*.

An external reflexivity with a crude internal pointing operates in consciousness of *level 2*. At this what appears in a perceptual field is determined in part by the perceptual powers of the perceiver. The contents of the field reflect to HIM what he is, but of course he does not find himself\* reflected in them. By hypothesis there is no *I* that could find anything.

Consciousness of *level 3* is a modification of consciousness of *level 2*. A central nucleus appears. It brings perspectival organization into the contents of the perceptual field. This perspective is a mixed, albeit blind, external-internal reflexivity. The perspective quietly reflects an orientation of the perceptual contents with respect to the perceiver. This orientation is certainly physical, especially in the case of visual perception. It is also psychological because the nucleus of the perceptual field exists thanks to the perceiver's attention. What is attended to reflects the attender's attention. Yet in the absence of *self-consciousness* the perspective is not apprehended as such.

Perceptual experience is primarily a phenomenon at *level 3*. One faces not merely a uniform perceptual field, but a focal and perspectival one. The perceptual judgments in which a perceptual experience culminates are judgments about the focal contents. These are demonstrative judgments about *this's* or *that's*, or *here's* or *there's* in the perceptual field in the *now* of the perceiving. Palpably, all of these items have an implicit reference to a potential, or actual, perceiving *I*.<sup>27</sup> The reference is implicit. As noted, it evinces a blind, mixed external-internal reflexivity.

At *level 4* an explicit reflexivity enters the stage. As this level has been described, the *I* here is just a blend of the reflexive sides of the contrasts *I-this/that* and *I-it* [external object of the world]. This *I* is a crude solipsistic one, which rounds up all the present experiences and unifies them as an *I* exhausted by them. This *I* collects and formalizes the internal/external axis.<sup>28</sup>

As characterized here *level 5* is a solipsistic consciousness. On the

chart offered, it is the last solipsistic level. As remarked above, the chart is merely a distillation of consciousness of *I*-strands; it says nothing about the causal independence, or dependence, or the chronological order of the strands. For all we know at this moment, perhaps solipsistic episodes of consciousness occur as the extreme cases within a non-solipsistic experience or life. Be this as it may, the execution of an act of will presupposes a consciousness in which an intended content of the form *I to do A [here] now* is brought into the causal process. This consciousness must not at the time of execution be concerned with other persons or objects. Its topic and concern are its effective causation. To be sure, other persons may still be involved in action A, as accusatives, beneficiaries, or circumstantial factors. However, action A need not involve any relation to the other: A may be a purely personal action for the exclusive benefit of the agent with no involvement of others. Hence, the consciousness of will, the acme of *level 5*, must always be not non-solipsistic. The *I* of deliberation is broader. The reasons for doing action A may involve all kinds of relationships to all sorts of persons. In principle, however, the deliberating *I* may believe that there exist no other *I*'s or persons. This belief may, of course, be a symptom of mental illness.

*Level 5* necessitates a special internal reflexivity. In deliberation the agent seeks to ascertain the range of his/her causal powers. Nonetheless, the search is not an intellectual aiming at a description of those powers. The search is a practical aiming at locating them, however blindly, in order to activate them. Hence, the agent's causal powers need be present in a *level 1* of consciousness.<sup>29</sup> Everything else is supernumerary.

At *level 6* the internality of *self*-reference becomes enhanced. Here a non-solipsistic *I*-strand appears. Yet the *I*'s with this *I*-strand as their social limit are isolated. They observe each other, but do not converse with one another. They can just contemplate their sharing of mental properties.

At *level 7* the internality of the reflexivity of *self*-reference grows by big strides. The full panoply of negativities determine the one instance that is apprehended in its total individuality. To reach this level of consciousness, a person must have resources for basking in personal relationships, for enjoying cooperative plans, and for experiencing personal conflicts.

Living requires that one rises to higher levels of consciousness when one encounters problems, and has to deliberate and then adopt plans of action. However, when one is engaged in carrying out those plans, one better not squander one's *self*-consciousness. Yet one need be attentive to the drift of events and be ready to make new decisions when obstacles turn up. Hence, the central level of consciousness for human living is *suffused level 3*, that is, suffused with the higher *I*-strands as potentialities for appropriate response.

V  
TAKING STOCK

In parts II–IV we have gone through a preliminary phenomeno-logical reflection on the reflexivity and self-reference of *self*-consciousness. This was in partial keeping with the promises made in part I. We have to say more about the doxastic pedestal of episodes of thinking. We need to explain in detail how indexicality is the categorial structure of subjectivity. This includes the thesis that the *I*'s are evanescent subjective particulars existing only within and during episodes of *self*-consciousness. (See note 9.)

We have seen over and over again how the contents of thought involve a contingent sameness, which must be thoroughly distinguished from strict identity. We have also seen that the worldliness/nonworldliness of the *I*'s and the other indexical particulars pivots on a contrast between sameness and strict identity. Hence, the results of our phenomeno-logical investigation need to be grounded on a theory of sameness. This theory must include an account of the contents of episodes of thinking as contrasted with doxastic attitudes. One theory that can do the job is Guise Theory. This is a comprehensive ontological account of the semantics of thinking. (See note 8.)

The analysis of the *I*-schemata so far developed is only a preliminary one. It needs to be enriched drastically before it can provide some guidance in the systematic study of *I*-disturbances.

Well, there is so much more to be done!

NOTES

1. For a theory of the sameness family of relations see Hector-Neri Castañeda, "Thinking and the Structure of the World," *Philosophia* 4 (1974): 4–40; "Identity and Sameness," *Philosophia* 5 (1975): 121–151—both in *Sprache und Erfahrung* (Frankfurt am Main: Suhrkamp, transl. Helmut Pape, 1982), the former in *Thinking, Language, and Experience* (Minneapolis: University of Minnesota Press, 1989), and in Klaus Jacobi and Helmut Pape, eds., *Das Denken und die Struktur der Welt* (Berlin: De Gruyter, 1989). On my view, the identity between mental and physical states is the sameness called in those papers *consubstantiation*. On the other hand, the basic sameness between veridical thought content and what makes it veridical hinges on the sameness *consociation*. Consubstantiation is also the sameness between the morning star and the evening star. Thus I seem to be endorsing a version of the Contingent-Identity Theory. See, for example, U. T. Place, "Is Consciousness a Brain Process?," *British Journal of Psychology* 47 (1956): 44–50; J. J. C. Smart, "Sensations and Brain Processes," *Philosophical Review* 68 (1959): 141–156; and David Lewis, "An Argument for the Identity Theory," *Journal of Philosophy* 63 (1966): 17–25, reprinted in his *Philosophical Papers, Vol. 1* (Oxford: Oxford University Press, 1983). My version of the contingent identity of mental events and processes to physical ones is not reductionist. See note 2. Furthermore, the theory is mounted on a general theory of contingent identity or sameness.

2. Many different types of mental fact constitute serious hurdles and tasks for reductionist programs. Here I mention just one fact seldom noticed. To me it has for the last thirty years seemed to be, not a mere hurdle, but a stumbling block for the reduction of episodes of consciousness to physical events. That fact is normal veridical, partially illusory visual experience. Not to generalize unduly, let me put the fact concretely. I am looking at the sky and see a triangle having as vertices the moon, the North Star, and the chimney on my house. This a veridical perception of the objects in question. The triangle is not a physical, but an illusory, one. Because of the time needed for light to travel, the position of the moon I see is one it occupied minutes before, whereas the position of the North Star is years older. My visual experience is precisely the visual presentation of what I see, that is, the mere existence of my visual field. My visual experience consists of the visual field containing the triangle just described. It does not occur in my brain (or whatever my thinking box may be). Of course, events in my brain have caused the existence of the presented visual field. My experience, however, occurs before my eyes, spread about in the piece of physical space containing the objects I see. My visual experience is *not* reductionally equivalent, let alone identical, to events in my brain (or thinking gadget). Because my veridical visual experience is partly illusory, the contents of my visual consciousness are not identical with the sub-domain of physical entities I see. (I say "entities" to include not so much physical objects in their fullness, but only their seen parts or surfaces.) My veridical visual displays occupy visual spaces which at least overlap with physical space. But even if such overlappings exhausted my visual spaces, there would be in visual space non-physical content. The physics and physiology of vision may causally explain why my visual spaces have certain non-physical, illusory contents. Nonetheless, causal explanation is not reduction. To cause is to bring into existence. (This is a basic truth often forgotten in defenses of reductionist causal claims.) Clearly, episodes of thinking, as well as dispositional states of believing, are individuated by their contents. In brief, my episodes of veridical visual consciousness are not reducible to complexes of physical events, including those within and those without my body. Patently, no theory of the world, or of the mind, or of consciousness can be satisfactory if it leaves my visual experience out of account. There is, however, a reductionism I have adopted: the economical view that reduces visual consciousness to the occurrence of visual content. For more data and details of these theses see Hector-Neri Castañeda, "Perception, Belief, and the Structure of Physical Objects and Consciousness," *Syntheses* 35 (1977): 285–351. This paper is an abridgment of part IV of *Sprache und Erfahrung*.
3. We have here no difficulties with Wittgenstein's arguments against private languages. See Ludwig Wittgenstein, *Philosophical Investigations* (Oxford: Blackwell, 1952), Elizabeth M. Anscombe, trans. We accept that our mature thinker has learned her language exactly as Wittgensteinians say: by having been engaged in social practices. We may also grant Wittgenstein's tenets that to follow a rule presupposes a community of rule-followers, and that internal psychological states must have external criteria.
4. Processes in the brain (or whatever) are the embodiments of thinking. Such processes must be sequences of events that embody representations of what is thought in a train of thinking episodes. The linguistic expressions of such thoughts have a characteristic grammatical form. Yet the embodying brain events need not be in a one-one correspondence with the spoken sentences or their components. What matters is that there be brain mechanisms that unfold those grammatical structures at the proper source of speech on the basis of the thinking-embodying events. Furthermore, the brain has to store mechanisms of representation, that is, mechanisms that have as outputs those thinking-embodying brain events. Likewise, the brain states or particulars that store such mechanisms need not involve the storing of items with the syntactic structure of the output sentences. In brief, whatever layers the hierarchy of brain mechanisms may have, at none of them is there a domain of brain items that is isomorphic to the sentences of the person's spoken language. The *syntacticism-connectionism* dispute is indifferent to our concerns here. These focus on the internal contents of the representations characteristic of, and peculiar to, self-consciousness.

5. Some distinguished philosophers hold this. See, for example, John Pollock, "My Brother, the Machine," *Nous* 22 (1988): 173–211; Roderick M. Chisholm, *The First Person* (Minneapolis: University of Minnesota Press, 1981); Sydney Shoemaker, "Self-reference and Self-awareness," *Journal of Philosophy* 65 (1968): 555–578.
6. Plato's life examination is not merely a quest for determining what one ought everything considered to do. He assumes that this *ought* is also a moral *ought*. He grounds his assumption on the view that human nature is such that immoral action is a manifestation of a sort of mental illness. The *Republic* is a subordinated, ambitious, and brilliant, yet bound to fail, effort at executing that grounding project.
7. The idea that what one thinks exists and what does not exist cannot have *any* properties at all is at least traceable to Parmenides's poem. Against him, Plato argued in the *Sophist* that non-existents are thinkable and can have some properties. Since then this issue has divided philosophers into antagonistic camps. For instance, William of Sherwood seemed to have adopted a Parmenidean position very much like Bertrand Russell's. See Klaus Jacobi, *Die Modalbegriffe in den logischen Schriften des Wilhelm von Shyreswood* (Leiden-Köln, E. J. Brill, 1980), 318ff. In this century Alexius Meinong and Bertrand Russell argued about the Parmenidean issue. Their focus was Meinong's existing round square. (For references and appraisal of this debate see my "Objects, Existence, and Reference: A Prolegomenon to Guise Theory," *Grazer Philosophische Studien* 25/25 (1985/86): 31–66, reprinted in Jacobi and Pape's *Das Denken und die Struktur der Welt*. It contains references to recent discussions of non-existing objects and of fiction.)
8. See the materials mentioned in note 1.
9. For an extensive study of indexical and quasi-indexical reference see Ch. 1, 4, and 12 of *Thinking, Language, and Experience*, or Ch. II.2, II.5, and II.6 of *Sprache und Erfahrung*.
10. On the difference and connections between indexical and corresponding quasi-indexical propositions see "Knowledge and Self: A Correspondence Between Robert M. Adams and Hector-Neri Castañeda" in Tomberlin's *Agent, Language, and the Structure of the World*: 293–309, reprinted in *Thinking, Language, and Experience*.
11. For more on the thinker's use of the first-person pronoun see "Response to Perry: Meaning, Belief, and Reference" in Tomberlin's *Agent, Language, and the Structure of the World*: 313–327.
12. Wilfrid Sellars's class lectures, 1950.
13. R. Descartes, *Oeuvres et lettres* (Paris: Bibliothèque de la Pleiade, 1941), André Bridoux, edit. 167.
14. Hector-Neri Castañeda, *Thinking and Doing* (Dordrecht: Reidel, 1975), Ch. 6, and *Thinking, Language, and Experience*, Ch. 7; and John Perry, "The Problem of the Essential Indexical," *Nous* 13 (1979): 3–21.
15. "Direct Reference, Realism, and Guise Theory: Constructive Reflections on David Kaplan's Theory of Reference" in Joseph Almog, John Perry, and Howard Wettstein, eds., *Themes of Kaplan* (Oxford: Oxford University Press, 1989).
16. See note 1. Guise Theory is essentially a theory of individuation and predication that distinguishes a rich family of sameness relations and conceives them as forms of predication. For discussion and developments of Guise Theory, see the exchanges with Romane Clark and Alvin Plantinga in Tomberlin's *Agent, Language, and the Structure of the World*; the exchanges with Jay Rosenberg, David W. Smith, Esa Saarinen, and Jeffrey Sicha in James E. Tomberlin, ed., *Hector-Neri Castañeda* (Dordrecht: Reidel Publishing Company, Profiles No. 6, 1986); the exchanges with Frank Döring, Heinz-Dieter Heckmann, Klaus Jacobi, Tomis Kapitan, Guido Küng, Wolfgang Künne, Anton Koch, Paolo Leonardi, Harald Pilot, Gerold Praus, Friedrich Rapp, Richard Schantz, in Jacobi and Pape's *Das Denken und die Struktur der Welt*.
17. See, for example, Roderick M. Chisholm, *Person and Object* (London: Allen & Unwin. 1976).

18. For a study of the family of the different types of negation and its contrasts with the family of colors properties, see Hector-Neri Castañeda, "Negations, Imperatives, Colors, Indexical Properties, Non-existence, and Russell's Paradox" in David F. Austin, *Philosophical Analysis: A Defense by Examples* (Dordrecht/Boston/London: Kluwer Academic Publishers, 1988).
19. Ferdinand de Saussure, *Cours de linguistique generale*, e.g., his view of plural as the whole network of contrasts, for instance, 'girl'/'girls', 'potato'/'potatoes', 'ox'/'oxen', etc.
20. Martin Buber, *I and Thou* (New York: Charles Scribner's Sons, 1970), trans. Walter Kaufmann.
21. See, for example, Hilary Putnam, "The Meaning of 'Meaning'" in Keith Gunderson, ed., *Language, Mind, and Knowledge* (Minneapolis: University of Minnesota Press, 1975), and "Meaning Holism and Epistemic Holism" in Konrad Cramer, Hans Friedrich Fulda, Rolf-Peter Horstmann, and Ulrich Pothast, eds., *Theorie der Subjektivität* (Frankfurt am Main: Suhrkamp, 1987); and Tyler Burge, "Individualism and the Mental" in Peter French, Theodore Uehling, and Howard Wettstein, eds., *Midwest Studies in Philosophy: Metaphysics* 12 (1979): 73–121. I am not convinced. I explain why in "Semantic Holism Without Semantic Socialism" forthcoming in *Midwest Studies in Philosophy: Philosophy of Language*.
22. This is a view that could be attributed to Kant as an interpretation of his celebrated slogan: "The *I think* must be able to accompany all my representations."
23. In "Perception, Belief, and the Structure of Physical Objects and Consciousness."
24. L. Weiskrantz, E. K. Warrington, M. D. Sanders, and J. Marshall, "Visual Capacity in the Menianopic Field Followed by Restricted Occipital Ablation," *Brain* 97 (1974): 709–728; L. Weiskrantz, "Trying to Bridge Some Neurophysiological Gaps Between Monkey and Man," *British Journal of Psychology* 68 (1977) 431–455; L. Weiskrantz, "Varieties of Residual Experience," *Quarterly Journal of Experimental Psychology* 32 (1980): 365–386.
25. P. S. Churchland, "A Perspective on Mind-Brain Research," *Journal of Philosophy* 77 (1980): 185–207.
26. John Pollock, in paper cited in note 5.
27. Chisholm attempts in *The First Person* to define 'this' and all other indicators in terms of 'I'. I argue for crucial differences between 'this' and 'I' in "Self-Consciousness, Demonstrative Reference, and the Self-Ascription View of Believing" in James E. Tomberlin, ed., *Philosophical Perspectives: I. Metaphysics* (Atascadero, CA: Ridgeview Publishing Co., 1987): 405–450.
28. See Hector-Neri Castañeda, "Consciousness and Behavior" in Hector-Neri Castañeda, ed., *Intentionality, Minds, and Perception* (Detroit: Wayne State University Press, 1967).
29. For a discussion of this see *Thinking and Doing*, Ch. 10, Sec. 3.