

THE FIRST-PERSON PERSPECTIVE: A TEST FOR NATURALISM

Lynne Rudder Baker

Self-consciousness, many philosophers agree, is essential to being a person. There is not so much agreement, however, about how to understand what self-consciousness is. Philosophers in the field of cognitive science tend to write off self-consciousness as unproblematic. According to such philosophers, the real difficulty for the cognitive scientist is phenomenal consciousness—the fact that we (and other organisms) have states that feel a certain way. If we had a grip on phenomenal consciousness, they think, self-consciousness could be easily handled by functionalist models. For example, recently Ned Block commented, “It is of course [phenomenal] consciousness rather than . . . self-consciousness that has seemed such a scientific mystery.” (Block, 1995, p. 230) And David Chalmers says that self-consciousness is one of those psychological states that “pose no deep metaphysical enigmas.” (Chalmers, 1996, p. 24) I think that this assumption that self-consciousness can be easily assimilated by science is too quick. For self-consciousness, as I shall try to show, rests on what I shall call “the first-person perspective.” And it is not obvious how to treat the first-person perspective scientifically.

In this paper, I shall describe the first-person perspective, and then argue that philosophers and cognitive scientists have neglected the first-person perspective at

their peril. At the end, I shall offer a challenge to naturalism: Either show how the first-person perspective can be understood naturalistically, or show that it is dispensable. My exploration of the first-person perspective here is “conceptual,” or at least pre-scientific. Although the term “self-consciousness” has been used to mean many things, all forms of self-consciousness presuppose the first-person perspective. I shall argue propaedeutically that there is no way that adequate accounts of various forms of self-consciousness can avoid the first-person perspective. My aim is not to convince you that the first-person perspective will forever elude science, but rather that no science that aspires to be a complete science of everything can afford to ignore it. Thus, the first-person perspective is a good test case for naturalism.

There are many preoccupations in the current literature that I am not going to address. For example, I am not concerned to define “consciousness.”¹ I am not concerned with whether or not all mental states are conscious.² Nor am I concerned with whether introspection can be conceived on a perceptual model.³ Nor am I concerned with whether or not there can be a science of consciousness.⁴ Nor am I concerned with the putative “explanatory gap” between processes in the brain and the ways that things taste, look, feel, smell, and sound.⁵ Nor am I concerned with the so-called

"hard problem" of consciousness: "Why is all this processing accompanied by an experienced mental life?" (Chalmers, 1996, xii) None of these issues, as topical as each is, is my interest here. Rather, I want to describe something that, in the welter of all the issues just mentioned, has been almost totally neglected in discussions by philosophers who see themselves as cognitive scientists, with unfortunate results for our understanding of human persons.⁶

TWO GRADES OF FIRST-PERSON PHENOMENA

A conscious being becomes self-conscious on acquiring a first-person perspective—a perspective from which one thinks of oneself as an individual facing a world, as a subject distinct from everything else.⁷ All sentient beings are subjects of experience (i.e., are conscious), but not all sentient beings have first-person concepts of themselves. Only those who do—those with first-person perspectives—are fully self-conscious. Beginning with nonhuman sentient beings, I shall distinguish two grades of first-person phenomena: weak and strong.

The weak grade of first-person phenomena is illustrated by problem-solving creatures whose behavior is explainable in terms of practical syllogisms—including how things seem to them. We attribute beliefs and desires (perhaps in the vocabulary of aversions, appetites and learning states) to nonhuman animals, who seem to be reasoning from a certain perspective. For example, the dog digs there because she saw you bury the bone there, and she wants it. (The fact that she stops digging when she finds the bone is evidence of the correctness of the explanation.) Or a researcher on infants hooks up a light so that it goes on and off when a neonate turns its head twice to the right; when the neonate figures out

how to control the light, she soon stops. (She gets bored.) And when the researcher changes the formula for turning on the light, the infant tries new combinations until she hits on the new formula. (Bower 1974; This example is discussed by Matthews 1985.)

Such explanations do not thereby attribute to the dog or to the infant any concept of itself as itself. Rather, they assume only that each organism has a certain perspective on its surroundings with itself as the "origin." The dog does not think of himself as himself or of himself as anything else; rather, we might say, the dog is the center of his own universe. He experiences things from his own egocentric perspective. If the dog could speak, he might say, from his own egocentric perspective, "There's a bone buried there in front of me, and I want it."⁸ Two points should be noted about weak first-person phenomena: (i) They are exhibited by sentient organisms, who solve problems by means of perspectival attitudes; these attitudes then explain the problem-solving behavior. (This point is independent of any theory of how, or whether, attitudes are explicitly represented in the brain.) (ii) No first-person concept is needed to bind belief-desire-behavior to a single organism. Since all the organism's psychological states are perspectival with the organism at the origin, the belief, desire and behavior all belong to the same individual, so to speak, by default. Although such an animal has beliefs and desires, he has no conception of belief or desire, nor of himself as the subject or bearer of beliefs and desires. He acts from his own perspective without any conception of having a perspective that differs from other perspectives.

The hallmark of a weak first-person phenomenon is that it is perspectival. (John Perry's work on indexicals illuminates this initial level of first-person phenomena. I

she soon stops. In the researcher turning on the combinations formula. (Bower discussed by

not thereby at the infant any. Rather, they as-ism has a certain dings with itself does not think of himself as any- at say, the dog is erse. He experi- own egocentric ould speak, he egocentric per- buried there in t."⁸ Two points ak first-person exhibited by sen- e problems by udes; these atti- problem-solving dependent of any er, attitudes are e brain.) (ii) No eded to bind be- single organism. s psychological ith the organism esire and behav- individual, so to gh such an ani- ires, he has no ire, nor of him- er of beliefs and own perspective aving a perspec- perspectives. first-person phe- spectival. (John illuminates this n phenomena. I

suspect that those who treat the first-person perspective so dismissively assume that all first-person phenomena can be understood as merely perspectival.⁹) In sum, animals whose behavior is explainable only in terms of their egocentric perspectives exhibit the weak grade of first-person phenomena.

Before turning to the strong grade of first-person phenomena, let me mention Gordon Gallup's well-known research with chimpanzees, who can be taught to recognize their bodies as their own. Chimpanzees' self-recognition seems to fall in between weak and strong first-person phenomena. Typically, when a nonhuman animal sees itself in a mirror, it responds as if it were seeing another animal. Gallup exposed young chimpanzees to a full-length mirror for 10 days. At first, they made other-directed responses (responses that would typically be made in the presence of other chimpanzees), but within 2 or 3 days, they began to display self-directed behavior—grooming parts of their body that were visually inaccessible without the mirrors, for example. Then Gallup anesthetized each animal and put red, odorless, nonirritating paint over one eyebrow and the opposite ear. After recovery from the anesthesia, the mirrors were reintroduced, and the chimpanzees began touching and exploring the marked areas at 25 times the rate before the mirrors were reintroduced. They smelled and visually examined the fingers that had touched the marked areas. Chimpanzees in the control group, who had not been exposed to mirrors before they were anesthetized and marked, displayed no mark-directed behaviors when mirrors were introduced. (With the exception of orangutans, other primates [such as monkeys] displayed no self-directed behavior after much longer exposure to mirrors.¹⁰) Gallup concluded that chimpanzees have a cognitive capacity that monkeys lack, and that those

chimpanzees who engage in self-recognition have rudimentary self-consciousness. (Gallup, 1977)¹¹

Another sort of intermediate case is illustrated by research on rhesus monkeys. One monkey ("O" for "operator") was placed in a divided box and taught to secure food by pulling one of two chains when a signal light was on. A second monkey ("SA" for "stimulus animal") was placed on the other side of the box, on a grid attached to an electric shock source, behind a one-way mirror, so that the O could see the SA, but the SA could not see the O. After three days during which the O adapted to the presence of the SA, the circuit was completed, so that when the O pulled one of the chains, the SA received a severe shock. Pulling the other chain when the light was on produced food but no shock to the SA. The experimenters varied the sequences and intervals of the light signals to ascertain the extent to which the O's pulling the shock-producing chain was influenced by the O's perception of the SA's agony. The conclusion was that "a majority of rhesus monkeys will consistently suffer hunger rather than secure food at the expense of electroshock to a conspecific."¹² One possible interpretation of the results is that rhesus monkeys can appreciate the points of view of conspecifics.

Without entering into the dispute over how to interpret Gallup's research or the research on the rhesus monkeys, I want to point out only that, on any reasonable interpretation, these chimpanzees and monkeys are still a long way from the self-consciousness enjoyed by humans.¹³ So, we must consider a higher grade of first-person phenomena.

A conscious being who exhibits strong first-person phenomena not only is able to recognize herself from a first-person point of view (as Gallup's chimpanzees did), but

also is able to think of herself as herself. For strong first-person phenomena, it is not enough to *distinguish* between first-person and third-person; also one must be able to *conceptualize* the distinction, to conceive of oneself as oneself. To be able to conceive of oneself as oneself is to be able to conceive of oneself independently of a name, or description or third-person demonstrative. It is to be able to conceptualize the distinction between oneself and everything else there is. It is not just to have thoughts expressible by means of "I," but also to conceive of oneself as the bearer of those thoughts. Nonhuman animals, who exhibit weak first-person phenomena, have subjective points of view. But merely having a perspective, or a subjective point of view, is not enough for strong first-person phenomena. Rather, one must also be able to conceive of oneself as having a perspective, or a subjective point of view.

Let me illustrate the distinction between weak and strong first-person phenomena in terms of grammar. Grammatically, we can distinguish between *making* a first-person reference (as when Smith says, "I am tall") and *attributing* first-person reference (as when Smith says, "Jones wishes that she (herself) were tall"). In the second case, Smith attributes to Jones a wish that Jones would express by saying "I wish that I were tall." The *attribution* of first-person reference occurs in indirect discourse, in a "that"-clause following a psychological (or linguistic) verb. However—and this is the important point—not only do we attribute first-person reference to others, but also we attribute first-person reference to ourselves—as when Jones says, "I wish that I were tall." A person who thinks, "I am tall" can distinguish herself from others; a person who thinks, "I wish that I were tall" can conceptualize that distinction, can think of herself as herself. The former

makes first-person reference; the latter *attributes* (as well as *makes*) first-person reference *to herself*. The ability to attribute to oneself first-person reference in indirect discourse ("I wish that I were tall") is a signal of strong first-person phenomena.

Following Hector-Neri Castaneda, I'll put an asterisk, a star, beside a pronoun to signal that it is an attribution of first-person reference.¹⁴ Call a sentence an "I*" sentence" if it is of this form: "I f that I* . . .," where "f" is replaced by a linguistic or psychological verb and "I* . . ." is replaced by a sentence containing a first-person reference. The use of an "I*" sentence—e.g., "I think (or hope or fear, etc.) that I* am F"¹⁵—is an indication that one is entertaining an "I*" thought. (But it is not an infallible indication; a computer could be programmed to produce "I*" sentences without having any thoughts at all.) An "I*" thought is one in which the thinker conceives of herself as herself*, without identifying herself by means of any third-person referential device, such as a name, description, or demonstrative. "I" is not a name for myself; I can use any name competently and still be mistaken about whose name it is. But I am never mistaken about who is picked out by my competent uses of "I." If I entertain an "I*" thought, I do not have to identify the person that I am thinking about; nor can I mistakenly believe that I am thinking about someone else. The ability to entertain "I*" thoughts—thoughts that attribute to oneself first-person reference in indirect discourse—is the ability to conceive of oneself as oneself*. I shall extend the use of the star ("*") to contexts in which the speaker is thinking of herself as herself* even in the absence of an embedded sentence.

In short, S can think of herself as herself* if and only if S can think of herself in a way naturally expressible in the grammatical first person as the bearer of

ence; the latter *at-*
akes) first-person
 ability to attribute
 reference in indirect
 ("I were tall") is a
 son phenomena.

Castaneda, I'll
 side a pronoun to
 ation of first-per-
 sentence an "I*"
 form: "If that I*
 ed by a linguistic
 "I*..." is replaced
 first-person refer-
 sentence—e.g.,
 etc.) that I* am
 at one is entertain-
 But it is not an
 computer could be
 "I*" sentences
 ights at all.) An
 which the thinker
 herself*, without
 ans of any third-
 such as a name,
 ative. "I" is not a
 se any name com-
 ken about whose
 r mistaken about
 y competent uses
 "I*" thought, I do
 person that I am
 mistakenly believe
 someone else. The
 oughs—thoughts
 t-person reference
 the ability to con-
 "I*. I shall extend
 contexts in which
 herself as herself*
 bedded sentence.
 f herself as her-
 think of herself
 ible in the gram-
 the bearer of

first-person thoughts. "I am tall" expresses a simple first-person thought. S can express her thought of herself as the bearer of the thought "I am tall" by saying, "I am having the thought that I* am tall." This latter sentence indicates that S is thinking of herself as herself*.

The ability to think of oneself as oneself* brings with it a number of related abilities. If one can think of oneself as the bearer of first-person thoughts, then one has the concept of a subject of thought and can think of others as subjects of thought. If one can think of oneself as oneself*, then, in addition to having desires (say), one can reflect on one's desires *as one's own*. (Conversely, without the ability to think of oneself as oneself*, one could not have the attitudes toward one's own desires ["second-order volitions"] that some take to be definitive of being a person. [See Frankfurt (1971).]) To be able to think of oneself as oneself is not just to have a perspective or subjective point of view—dogs have perspectives—but also to be able to think of one's perspective as one's own, and to think of others as having different subjective points of view from one's own.

Anyone who has the ability to conceive of herself as herself* has the ability to conceive of other things as different from herself. From a conceptual point of view, it seems that one cannot conceive of oneself as oneself* unless one can conceive of oneself as distinct from other things. And one cannot conceive of oneself as distinct from other things unless one has concepts of things as other. (The ability to conceive of things as distinct from oneself is required in order to doubt that such things exist.) This conceptual point is borne out by—or rather taken for granted by—developmental psychologists, who routinely describe the acquisition of self-concepts in tandem with the acquisition of

concepts of other things as different from oneself. (Stern, 1985)

An ability to conceive of oneself as oneself* in the sense just described is both necessary and sufficient for the strong grade of first-person phenomena: An individual who is the locus of strong first-person phenomena can conceive of herself as a bearer of first-person thoughts. She manifests an ability not only to make first-person reference, but also to attribute to herself first-person reference. Strong first-person phenomena do not seem reducible to purely non-first-person phenomena. For example, the thought expressed by "I regret that I* had to be the one to break the news of the nonrenewal of your fellowship" is not adequately paraphrased as "LB regrets that LB had to be the one to break the news . . ." (which substitutes attribution of a third-person reference for the attribution of first-person reference). You may adequately report what I said by "LB regrets that s/he* had to be the one to break the news . . .," but that sentence retains the attribution of first-person reference to me by the speaker. If I attribute first-person reference to myself, my sentence cannot be adequately paraphrased by any sentence that fails to attribute first-person reference to me: The attribution of first-person reference to oneself seems to be ineliminable. (Baker 1981)

We can sum up the grades of first-person phenomena this way: Weak first-person phenomena are exhibited by problem-solving beings whose behavior is explained by attitudes understood perspectively, from their own points of view. If a dog that exhibits weak first-person phenomena could express its attitudes in English, it would locate things relative to its own spatio-temporal position (e.g., "There's danger over there"). But it would not thereby show that it had any concept of itself or even any

ability to recognize itself from a first-person point of view. It simply acts from its own perspective, with itself as the center. All experience of any sentient being is perspectival, had from its own point of view. It is characteristic of weak first-person phenomena that they are perspectival in this way. On the other hand, strong first-person phenomena require that the subject conceptualize the distinction between herself from a third-person point of view and herself from a first-person point of view. The subject of strong first-person phenomena is not only able to think first-person thoughts (typically, using "I"), but is also able to attribute to herself first-person thoughts (typically, using "I*"). Not only can she think of herself, but she can think of herself as herself*, and of her thoughts as her own*.

Those with the ability to exhibit strong first-person phenomena have a first-person perspective. Although Gallup's chimpanzees have some claim to making first-person reference, and the rhesus monkeys in the study described earlier appear to appreciate the perspectives of conspecifics, I shall reserve the term "first-person perspective" for the subjects of strong first-person phenomena. One has a first-person perspective, then, if and only if one has the ability to conceive of oneself as oneself*, where this ability is signaled by the linguistic ability to attribute (as well as to make) first-person reference to oneself.

In short, to have a first-person perspective is to have a certain *ability*. This ability is manifested typically, but not only, by the use of "I*" sentences. It is also manifested on occasions when one has a thought that one would express by a non-"I*" sentence—e.g., "Why, it's *me* they're talking about!" To take another example, one manifests a first-person perspective by wondering, "Am I the winner?" For one

could not ask this question if she lacked the conceptual resources to think, "I wonder whether I* am the winner." So, although a sentence like, "Am I going to die, Doc?" is not an "I*" sentence, it does indicate that the questioner has a first-person perspective. A first-person perspective is manifested any time that a person has a thought, however it is expressed, that could not be entertained by anyone who lacked the ability to think of oneself as oneself*.

Although, from a first-person perspective, I have the ability to think of myself in a unique way, there is no funny object that is myself-as-myself; there is no entity that is "self"—other than the person who I am.¹⁶ The referent of "I" and of "I*" is the person: not a body, not a disembodied ego. "I*" does not denote a spooky entity to which I alone have direct access, "self." When I say, "I wonder whether I* will still dream of being unprepared for class in ten years," I refer twice to myself—to the person, LB, in my embodied concreteness. When I refer to myself by means of "I*," *what* I refer to is no different from what you refer to by means of "LB." What is special about "I*" is that I can conceive of that person in a way that you cannot, from "the inside," so to speak. Descartes's discovery—or rather his rediscovery, after Augustine—of "the inner" is the real contribution of *Meditations* II. (cf. Matthews 1992). It is Descartes's reification of inwardness that I reject: what is distinctive about being a person does not need to be secured by a logically private entity, to which no one has access but me.

So, "I" and "I*" have the same referent. Do they also express the same concept? Assuming that "I" and "I*" express concepts at all, then use of "I" by a toddler (who does not yet exhibit strong first-person phenomena) does not express the full self-concept that he later expresses with both "I" and "I*" after he

on if she lacked to think, "I won- winner." So, "Am I going to sentence, it does er has a first-per- person perspective at a person has a essed, that could one who lacked self as oneself*.

person perspec- think of myself no funny object there is no entity the person who I and of "I*" is the disembodied ego. spooky entity to access, "self." ether I* will still I for class in ten self—to the per- d concreteness. means of "I*," rent from what "LB." What is can conceive of ou cannot, from Descartes's dis- discovery, after is the real con- l. (cf. Matthews ification of in- at is distinctive not need to be ivate entity, to ut me.

ame referent. Do ncept? Assuming epts at all, then does not yet ex- ena) does not pt that he later d "I*" after he

has acquired a full first-person perspective. For a being without a concept of itself as itself*, "I" is just a marker of perspective. Acquisition of a first-person perspective brings with it a genuine self-concept—a concept of oneself as oneself*. So, for a being who has come to conceive of itself as itself, "I" and "I*" express the same self-concept. But before acquiring a first-person perspective, a being who uses "I" does not have that full self-concept. Such a child has not fully mastered the use of "I." Complete mastery of "I" includes the ability to use "I*."

The relation between the first-person perspective and self-consciousness is this: The first-person perspective is a necessary condition for any form of self-consciousness, and a sufficient condition for one form of self-consciousness as well. For a conscious being with a first-person perspective can conceive of her thoughts, attitudes, feelings and sensations as her own*. And the ability to conceive of one's thoughts and so on as one's own* is a form of genuine self-consciousness. Every other form of self-consciousness that I know of presupposes self-consciousness in this basic sense. So, every self-conscious being—"self-conscious" in any sense whatever—has a first-person perspective.

THE RELATIONAL CHARACTER OF THE FIRST-PERSON PERSPECTIVE

There are several unusual features of the first-person perspective. The fact that first-person reference is immune to a certain kind of referential error, for example, has been amply discussed.¹⁷ Here I want to call attention to a different, and perhaps more controversial, feature of the first-person perspective. A first-person perspective is relational in that it would be impossible for a being truly alone in the universe to have a first-person perspective. One cannot think of oneself as oneself* without concepts of

other things by means of which to distinguish things as being different from oneself; and one cannot have concepts of other things without the presence of other things. (For an elaboration of this theme, see Sartre 1966.) It is only over and against other things in the world that one stands as subject with a first-person perspective.¹⁸ Here is a simple argument:

Argument for the Relational Character of the First-Person Perspective

- (1) x has a first-person perspective if and only if x can think of herself as herself*.
- (2) x can think of herself as herself* only if x has concepts that can apply to things different from x.
- (3) x has concepts that can apply to things different from x only if x has had interactions with things different from x.

Therefore,

- (4) If x has a first-person perspective, then x has had interactions with things different from x.

So, x's having a first-person perspective depends upon x's relations to other things. Therefore, the property of having a first-person perspective is, in one sense, a relational property. The claim is that x has a first-person perspective only if x is embedded in a world of things from which x can distinguish herself as herself*. The argument is valid, but is it sound?

Premise (1) is simply a restatement of the definition of a first-person perspective. Premise (2) is what I argued earlier is a conceptual truth, taken for granted by developmental psychologists. The controversial premise is the anti-Cartesian premise (3). (3) is a strong rejection of "internalism" in the philosophy of mind, against which I have argued at length elsewhere. (Baker 1987, pp. 23–105; Baker 1995, pp. 42–56) Here I'll simply point out the consequence

of denying (3). The main consequence is that denial of (3) leaves one with no plausible account of concept-acquisition. Descartes, who would deny (3), never questioned whether his own use of concepts in his *Meditations* was at odds with his ontological assumptions. He just assumed that he could bracket his beliefs about things other than himself and still have available his empirical concepts of other things (e.g., "sitting in front of the fire") as well as the concepts required to think, "I am certain that I* exist." Descartes did not ask himself how, if he were the only finite being, he could have acquired the concepts needed to have all these thoughts.

Perhaps, a Cartesian may say, Descartes acquired a concept of "sitting" and of "in front of" and of "fire" from the fact that it *seemed to him* that he is sitting in front of a fire. But this is a nonstarter: If Descartes could reason from the premise that it seemed to him that he was sitting in front of a fire, then he must *already* have had the concept of sitting in front of a fire, and hence the concepts of "sitting" and of "in front of" and of "fire." But it is the acquisition of these concepts (of "sitting" and of "in front of" and of "fire") that we are wondering about.

Perhaps, a Cartesian may say, all concepts are innate; in that case, Descartes was born with concepts of "sitting" and of "in front of" and of "fire." But this cannot be right: The innateness hypothesis cannot be taken in so strong a sense. For in this overly strong sense, everybody not only is born with a concept of "sitting" and so on, but also with the concept of "quarks." But that's false; Descartes was a genius, but he had no concept of quarks. Moreover, the sense (if any) in which Descartes was born with a concept of sitting in front of a fire is not a sense in which it would be available for his use in reasoning. (He could not

reason about sitting in front of the fire until years after his birth.) However, it seems partly definitive of having a concept—whether concepts are mental representations, skills or something else—that it be available for use in reasoning. So, we cannot say that Descartes was born with a concept of sitting in front of a fire (in the relevant sense).¹⁹

Perhaps, a Cartesian would say, Descartes was born with concept-seeds, so to speak, that grew in him (like toenails) without any interactions with finite things other than himself. Or perhaps a Cartesian would say that God put all the relevant concepts in Descartes's mind at the appropriate time. I find such suggestions utterly implausible. A Cartesian may object: "Sure," he may say, "these suggestions are empirically implausible, but, still they are metaphysically possible." I am at a loss about how to respond—except to say this: In virtue of what does someone have a concept of *fire* independently of his interactions with other things (maybe not fires, but lighted cigarettes or electric stoves)? I'm not asking for evidence; I'm asking what would make it the case that, independently of his interactions with other things, a person has one concept and not another? A particular state of a brain or a soul, or a particular mental representation? But the original question won't stay down: What makes it the case, independently of interactions with the environment, that a particular state of a brain or a soul is a concept of "sitting" or of "in front of" or of "fire," or that a particular mental representation represents any of those things? These questions concern not just contingent facts about the acquisition of concepts, but rather (noncontingent) facts about what it is to have empirical concepts. And I see no possible answers to such questions, without adverting to the environment.

of the fire un-
 however, it seems
 a concept—
 representations,
 that it be avail-
 So, we cannot
 with a con-
 a fire (in the

id say, Descartes
 ds, so to speak,
 ils) without any
 ings other than
 esian would say
 ant concepts in
 appropriate time. I
 ily implausible.
 "Sure," he may
 re empirically
 are metaphysi-
 ss about how to
 is: In virtue of
 concept of *fire*
 tions with other
 ut lighted ciga-
 I'm not asking
 hat would make
 ntly of his inter-
 person has one
 particular state
 particular mental
 iginal question
 akes it the case,
 ons with the en-
 state of a brain
 itting" or of "in
 hat a particular
 presents any of
 ons concern not
 the acquisition
 (noncontingent)
 have empirical
 ible answers to
 dverting to the

Perhaps a contemporary Cartesian would even agree that one must also stand in a certain causal relation to certain things in the external world (like fire), and add that in order to have the concept of fire, one must be prepared to make certain inferences from certain sorts of sense experience. But the Cartesian may continue, in addition to the concept of fire, we may imagine an alternative concept, "fire₂"—a concept that has the same inferential role as "fire," but which can be possessed without satisfying any externalist causal requirement. In that case, it may be claimed that Descartes did have the concept "fire₂," and, furthermore, that concepts like "fire₂" are the only kind of concept needed to satisfy (2) of the Argument for the Relational Character of the First-Person Perspective. So, the contemporary Cartesian may conclude, premise (3) is false and the argument is unsound.

Now premise (3) is false only if a person can acquire the concept "fire₂" without any interaction with things different from herself. But the same old question comes up: Under what conditions would a person have the concept "fire₂"? The contemporary Cartesian cannot just help himself to the notion of inferential role.²⁰ For if a person were the sole occupant of the universe, what would count as inferring that *p* rather than inferring that *q*? In virtue of what would something be an inference at all? The contemporary Cartesian cannot advert to its *seeming* to the person that she is inferring that *p*, or inferring anything at all. For not only is it the case that someone's inferring (as opposed to its seeming to her that she is inferring) does not require that she have the concept of inferring, but also its *seeming* to someone that she is inferring does require that she have the concept of inferring, a concept of whose acquisition we have no Cartesian account. Nor can the contemporary Cartesian advert to the

tokening of Mentalese symbols (or to brain states). For, by hypothesis, *none* of these Mentalese symbols (if there are any) is interpreted. The concept of "fire₂" is not like the idea of Hume's missing shade of blue, for which a place has been made, so to speak by impressions of other colors. If a person were alone in the universe, it would be an implacable mystery what would count as having one concept rather than another.²¹ So, pretending that there are alternative concepts like "fire₂" provides no good reason to deny (3).

I do not claim to have refuted those who deny (3), but I do hope to have made plain the difficulty of holding this aspect of Cartesianism. Although I have no theory—Cartesian or non-Cartesian—of concepts or of concept-acquisition, I would look to Wittgenstein for the former and to developmental psychologists for the latter. Both sources—Wittgenstein and developmental psychologists—rely heavily on interactions with the environment as partly determinative of one's concepts.

It is noteworthy that (3) completely defuses the threat of solipsism, understood as the view that nothing but me exists. For (3) shifts the issue of solipsism from the question often asked of beginning students, "What reason is there to reject solipsism?" to a prior question: "Is solipsism conceivably true?" If I can formulate the thesis of solipsism, then I have concepts applicable to other things besides myself. And if I have concepts applicable to other things besides myself, then (given [3]) solipsism is conceptually false; and if solipsism is conceptually false, then no further reason is required to reject it. The Argument for the Relational Character of the First-Person Perspective implies that Descartes could not even have raised his skeptical question unless he had already been guaranteed a nonskeptical answer.

THE INDISPENSABILITY OF THE FIRST-PERSON PERSPECTIVE

Two different sorts of considerations suggest that the first-person perspective is indispensable for our theorizing about reality. The first sort (I) concerns language: First-person reference is not eliminable from "I*" sentences, whether it is eliminable from simple, direct-discourse "I" sentences or not. The second sort (II) concerns psychological explanation: Certain psychological explanations of behavior require attribution of a first-person perspective to the one whose behavior is to be explained.

(I) Although the first-person perspective does not depend on natural language, it is often manifested, as we have seen, in a person's use of "I*" sentences. I want to show that "I*" sentences differ from simple, direct-discourse "I" sentences in that first-person reference is ineliminable from "I*" sentences, whether it is eliminable from simple "I" sentences or not. Consider first-person contexts in which some philosophers have suggested that we might do without the grammatical first person. According to Peter Geach, for example, the pronoun "I" serves to call attention to the speaker, and in soliloquy, "I" is "idle, superfluous." If Descartes had said, "I am getting into an awful muddle," Geach said, he could have expressed himself equally well by saying, "This is an awful muddle."²² Bertrand Russell made a similar point: The premise "I think" could be rephrased as "There is thinking," (and hence, said Russell, does not support the conclusion "I exist," where "I" refers to a substantial self). (Russell, 1956, p. 567)

Russell and Geach give reason to conjecture that use of "I" in simple sentences, together with an inability to use "I*" sentences, would be an indication of weak first-person phenomena. Imagine that a

dog, who lacks a first-person perspective, could talk. As a subject of weak first-person phenomena, however, the imagined dog may well utter simple first-person sentences like "I see a potential mate." Such an utterance could as well be rendered as "There's a potential mate over there." But lacking a first-person perspective, the dog would not have the ability meaningfully to utter, "I hope that I* will find a suitable mate." On the basis of the remarks of Russell and Geach, I conjecture that the grammatical first-person could be eliminated altogether for beings that enjoyed weak first-person phenomena, but who lacked a full first-person perspective.

The situation is rather different for beings who have a full first-person perspective and who speak a language like English. For English speakers with first-person perspectives have the capacity meaningfully to utter "I*" sentences, from which first-person reference is ineliminable. The thought that one expresses by "I*" could not be equally well expressed in a non-first-person way.²³ For example, there is no third-person way to express the Cartesian thought, "I am certain that I* exist." The certainty that Descartes claimed was certainty that *he** existed, not certainty that *Descartes* existed. And these states of certainty are not equivalent. (Castaneda 1966, 1967; Baker 1981) So, "I am certain that I* think," which indicates a first-person perspective, cannot be subjected to the same treatment that Russell proposed for "I think," which plausibly may be rendered, "There is thinking." The "I*" is ineliminable. In sum: Whether or not Geach and Russell were right about the eliminability of "I" in direct discourse, their point would not apply to the use of "I" in indirect discourse. First-person reference is ineliminable when "I" is used as "I*" in indirect discourse, where such use indicates a first-person perspective.

erson perspective, of weak first-personer, the imagined first-person sentimental mate." Such will be rendered as "over there." But perspective, the dog meaningfully to find a suitable the remarks of conjecture that the could be eliminating that enjoyed phenomena, but who perspective.

fferent for beings perspective and like English. For first-person perspective meaningfully to which first-personable. The thought "I" could not be a non-first-personable, there is no the Cartesian "I* exist." The claimed was certainty that states of certainty that Castaneda 1966, am certain that is a first-person subjected to the all proposed for may be rendered." The "I*" is whether or not right about the correct discourse, to the use of "I" person reference used as "I*" in such use indicative.

(II) The second way in which the first-person perspective may be seen to be indispensable is in psychological explanation. Some psychological explanations require a first-person perspective in that they attribute attitudes that would be unavailable to an agent who lacked a first-person perspective. For example, part of the explanation of Oedipus's blinding himself is that he came to realize that he* was the killer of Laius. Oedipus may have expressed this realization without using an "I*" sentence, by saying for example simply, "I killed Laius." But the psychological state that helps explain his self-blinding—the realizing that he* killed Laius—requires that Oedipus have the ability to think of himself as himself*.

We can see this in two ways. (i) Oedipus's realization that he* was the killer of Laius included an understanding that earlier, when he had been looking for the killer, he had not realized that he* was the killer of Laius. We have seen that one's wondering whether she* is the winner requires that she have the conceptual resources to entertain the thought expressible as "I wonder whether I* am the winner." In the same way, Oedipus's coming to realize that he* is the killer of Laius required that he have the conceptual resources to entertain the thought expressible as "Although I did not realize it before, I now realize that I* am the killer of Laius." No one lacking a first-person perspective would have the ability to have that thought. So, whether actually expressed in an "I*" sentence or not, Oedipus's realization was one that required that he have the ability to conceive of himself as himself*.

(ii) On one reasonable interpretation, in order for Oedipus's realization that he* killed Laius to have motivated him to blind himself, Oedipus had to be thinking of himself as himself*. I am not just remaking

John Perry's point that action is explained by belief states understood indexically.²⁴ Indexically characterized belief states (sometimes called "self-locating beliefs") situate an agent perspectivally in an environment. That self-locating ability, as we saw earlier, is shared by all problem-solving animals (who are subjects of weak first-person phenomena), and does not suffice for a first-person perspective. Nor do mere self-locating beliefs explain Oedipus's blinding himself. For example, the following line of reasoning, which includes self-locating beliefs, would not suffice as an adequate explanation of Oedipus's self-blinding: "Whoever killed Laius should be blinded; I killed Laius; therefore, I should be blinded." Rather, Oedipus blinded himself because of the horror of the realization that *he himself* had killed Laius. Nothing less than a first-person perspective would do justice to Oedipus's motivation. If this is right, then in order to explain Oedipus's blinding himself, we must suppose that he not only was subject of weak first-person phenomena that would enable him to have first-person self-locating beliefs, but also that he had a conception of himself from a first-person point of view. Since the correct psychological explanation of Oedipus's blinding himself requires that Oedipus had the ability to conceive of himself as himself*, a psychology that aims to be a complete theory of behavior cannot afford to ignore the first-person perspective.

A LOOK AT OTHER VIEWS

The first-person perspective is the key to self-consciousness. Without trying to offer an analysis or theory of self-consciousness beyond the characterization that I have already given, I now want to show that the first-person perspective causes trouble for recent prominent views on self-consciousness. To see what is at stake, first

consider two ways that David Rosenthal's "higher-order thought" account might be interpreted. Rosenthal builds up what he calls "introspective consciousness" by a kind of iteration: A "mental state is conscious—non-introspectively conscious—just in case one has a roughly contemporaneous thought to the effect that one is in that very mental state." And the mental state is "introspectively conscious" when the relevant contemporaneous second-order thought is itself conscious. "Since a state's being conscious is its being accompanied by a suitable higher-order thought, introspective consciousness occurs when a mental state is accompanied both by such a second-order thought, and also by a yet higher-order thought that one has that second-order thought." (Rosenthal 1993, p. 199) To report that one is in a certain mental state is to express a higher-order thought in virtue of which the first-order mental state is conscious. If Sally says, "I have a headache," she is reporting her headache by expressing a higher-order thought in virtue of which she is conscious of her headache.

Rosenthal puts it this way: "[I]n general, our being conscious of something is just a matter of our having a thought of some sort about it. Accordingly, it is natural to identify a mental state's being conscious with one's having a roughly contemporaneous thought that one is in that mental state." (Rosenthal 1991, p. 465) He makes a similar point when he says that "a mental state's being conscious will be the same as one's having the ability to express one's higher-order thought that one is in that mental state." (Rosenthal 1993, p. 204) How are we to understand "one" in the "higher-order thought that one is in that mental state"? Is it one* (oneself as oneself*) or just oneself? Thoughts that simply happens to be about oneself—as Oedipus's thoughts about the killer of Laius before his awful

realization—do not require that one have a first-person perspective; thoughts about oneself as oneself* do require that one have a first-person perspective. So the question—How are we to understand "one" in the "higher-order thought that one is in that mental state"?—gives rise to two ways of formulating this account of consciousness in terms of higher-order thought. The first interpretation leaves out the first-person perspective; the second interpretation—Rosenthal's own²⁵—requires a first-person perspective. The point that I want to emphasize by giving two interpretations of Rosenthal's view is the importance of the first-person perspective.

First interpretation: Mental state *M* is introspectively conscious if and only if there is some thinker or subject *S* and some time *t* such that:

1. At *t*, *S* is in mental state *M*.
2. At *t*, *S* has a thought that *S* is in mental state *M*.
3. At *t*, *S* has a thought that *S* has a thought that *S* is in mental state *M*,

where *S*'s reporting the thought that she had in (2) is sufficient for the truth of (3).

Second interpretation: Mental state *M* is conscious if and only if there is some thinker or subject *S* and some time *t* such that:

1. At *t*, *S* is in mental state *M*.
- 2a. At *t*, *S* has a thought that she* is in mental state *M*,
- 3a. At *t*, *S* has a thought that she* has a thought that she* is in mental state *M*,

where "she*" attributes a first-person reference to *S* and where *S*'s reporting the thought that she has in (2a) is sufficient for the truth of (3a). As we have seen at length, sentences like (3a), which attribute first-person reference, cannot be replaced without loss by any sentences—like (1) or iterations of (1)—that are free of the first person.

that one have thoughts about that one have. So the question—stand “one” in that one is in that two ways of consciousness right. The first the first-person interpretation—a first-person I want to emphasize interpretations of importance of the

state M is in—only if there and some time t

S is in mental

has a thought

right that she the truth of (3).

mental state M is some thinker such that:

he* is in men-

at she* has a mental state M,

st-person re-reporting the

is sufficient

have seen at

which attribute

be replaced

—like (1) or

of the first

Let us begin with the first interpretation, which leaves out the first-person perspective altogether. On the first interpretation, the “higher-order-thought” account is subject to counterexamples. Satisfaction of conditions (1)–(3) does not account for a thinker’s being in a mental state of which she is introspectively conscious. Suppose that Jones is being tested on her ability to read PET scans, and that she is reading a contemporaneous PET scan of her own brain. Although she knows that she is reading a PET scan of an alert subject named “Jones,” she does not realize that that Jones is herself*. (She thinks that the Jones whose brain she is watching is in the next room.) Now suppose that the telephone rings in the next room, where she thinks that the subject Jones is located. At *t*, she points to a lit-up portion of the brain on the screen and says to the tester, “Now Jones is hearing the phone.” In so saying, Jones is expressing her thought that now Jones is hearing the phone. At the same time, with no conscious inference, Jones thinks to herself, “Jones is having the thought that Jones is hearing the phone.”²⁶

This story, I believe, satisfies conditions (1)–(3) for Jones to be introspectively conscious of hearing the phone at *t*: Jones hears the phone at *t*, thus satisfying (1); Jones has a thought that Jones hears the phone at *t*, thus satisfying (2); and Jones reports that thought, thereby expressing the third-order thought that Jones has a thought that Jones is hearing that phone and thus satisfying (3). Yet, in the circumstances described, satisfaction of (1)–(3) does not make Jones introspectively conscious of hearing the phone at *t*. Even if we stipulate that Jones is indeed introspectively conscious of hearing the phone at *t*, it is not in virtue of satisfying (1)–(3) that Jones is introspectively conscious of the hearing of the phone at *t*. For Jones (mistakenly)

believes that the person whose hearing of the phone she is talking about is someone other than herself*. In short, Jones’s thought that Jones is having the thought that Jones is hearing the phone does not make Jones’ hearing the phone introspectively conscious.²⁷ So, omission of the first-person perspective by the first interpretation of the “higher-order-thought” account, (1)–(3), leads to counterexamples. No non-first-person thought could possibly confer introspective consciousness on a mental state.

Whether or not the second interpretation—(1), (2a), and (3a)—ultimately provides a satisfactory account of introspective consciousness, it does build in the first-person perspective. The issue of interest here, however, is that the second interpretation, whether otherwise satisfactory or not, gives no purchase on a reductive explanation of consciousness. Rosenthal hopes to show “how consciousness can occur in physical things” by “explaining the consciousness of mental states in terms of mental states that are not conscious.” (Rosenthal, 1991, p. 474) So, on the second interpretation, the explanation of “the consciousness of mental states in terms of mental states that are not conscious” is that a first-order mental state becomes conscious by being accompanied by a second-order (nonconscious) state. And the accompanying second-order state itself becomes conscious when it is accompanied by a (nonconscious) third-order state that *only a being with a first-person perspective can have*.

This explanation, as it stands, I think, cannot be a fully reductive explanation of consciousness. For the third-order nonconscious states that are to explain introspectively conscious states cannot be understood simply in terms of less complex states; rather, the third-order explanatory

states themselves require what the first-order mental states do not—namely, a first-person perspective. So, on Rosenthal's account as developed so far, the first-person perspective would itself remain unexplained and unreduced.

To see this, consider again the mental state of hearing the phone. On the second interpretation of Rosenthal's view, this state becomes conscious when I have the second-order thought expressed by

(a) I am hearing the phone;

and the state of hearing the phone becomes introspectively conscious when I have the third-order thought expressed by

(b) I am having the thought that I* am hearing the phone.

It may seem that the mental state expressed by (b) can be understood as a mere iteration of the mental state expressed by (a).²⁸ But, as we have seen, not every being with the capacity to entertain the thought expressed by (a) has the capacity to entertain the thought expressed by (b). All that is required to entertain the thought expressed by (a) is the kind of perspectival consciousness that we attribute to dogs and other problem-solving creatures who are subject to weak first-person phenomena. But in order to entertain the thought expressed by (b), a being must have a first-person perspective—something that a dog lacks. So, the thought expressed by (b) introduces a new factor—a first-person perspective—that is absent from the lower-order states.

If an explanatory account requires a new factor that is absent from the states to be explained, then the account is not reductive unless the concept of the new factor itself is either part of a reductive apparatus (as, e.g., the concept of molecules is) or the new factor is itself reduced (as, e.g., heat is reduced to molecular motion). The concept of the first-person perspective is

not part of a reductive apparatus; so, unless the first-person perspective can be reductively explained, no account in which it figures is fully reductive. This is so even if Rosenthal's account does succeed in (reductively) explaining conscious states in terms of nonconscious states. My point is twofold: (a) the nonconscious states in terms of which the conscious states are to be explained introduce something new, the first-person perspective, and, therefore, (b) the account is not fully reductive unless the first-person perspective itself is reductively explained.²⁹

At best, the second interpretation of Rosenthal's view reduces introspectively conscious states to nonconscious states that require a first-person perspective, with no attempt to reduce the first-person perspective. For this reason, I do not think that Rosenthal's view, even if successful as an account of conscious states, provides the kind of reduction that robust naturalists hanker after. For the first-person perspective that is required for the explanation of conscious states is itself left unexplained.

The upshot of this discussion of the "higher-order-thought" account of introspectively conscious states is a dilemma for the naturalizer of self-consciousness: On the one hand, if (as on the first interpretation) the account is indifferent to the distinction between conceiving of oneself as oneself* and simply having a point of view that allows one to conceive of oneself in a third-person way, then it leaves out the first-person perspective; and the account is subject to counterexamples. On the other hand, if (as on the second interpretation) the account invokes the first-person perspective, then it does not succeed in reductively explaining consciousness or self-consciousness; and is incumbent upon the naturalizer to give a naturalistic account of the distinction in question and our ability to make it.³⁰

Let me briefly mention three other prominent approaches to self-consciousness, and show how they too falter over the first-person perspective. Consider the accounts of self-consciousness: as self-scanning of internal states, as approachable from a third-personal point of view, and as possession of a "narrative self."

(A) Some philosophers, such as D. M. Armstrong, take self-consciousness (or introspection) to be "a self-scanning process in the brain." (Armstrong 1968, p. 324; Lycan 1987, pp. 72–3; cf. Lycan 1996) Perhaps, but self-scanning as it is now understood is not sufficient for self-consciousness. A self-scanner cannot make the distinction that is crucial for self-consciousness: A self-scanner cannot distinguish between acquiring information about something-that-is-in-fact-itself and acquiring information about itself-as-itself. Suppose that S is a system that has a scanner S' of its internal states. S' acquires information about S and feeds it to a device that regulates S's states. In a sense, then, S acquires information about itself. But in the system, there is no place for a distinction between saying, "S acquires information about S" and saying, "S acquires information about itself as itself." And the *use* made of S's information about S's internal states could be described as well by saying that S regulates S's states as by saying that S regulates its own states. So, the behavioral difference that the first-person perspective makes for self-conscious beings is wholly absent in the case of self-scanning scanners. Since S cannot distinguish between itself-as-S (from the third person) and itself-as-itself, the system makes no distinction between scanning S and scanning itself. And without that distinction, there is no self-consciousness in S. Therefore, self-scanning scanners do not account for self-consciousness.

Let me be more cautious: Perhaps in the future self-scanning scanners will be able to accommodate the crucial distinction between a first-person perspective and a third-person perspective on oneself. I just can't imagine how. The point I want to make here, however, is that writers on the mind have not even tried to accommodate the crucial distinction, and until they do, they have not addressed a feature essential to self-consciousness.

(B) Some philosophers explicitly eschew first-person approaches to consciousness and to self-consciousness. Daniel Dennett, for example, is emphatic: A theory of consciousness "will have to be constructed from the third-person point of view," he says, "since *all* science is constructed from that perspective." (Dennett 1991, p. 71) Despite his official view, however, Dennett's practice takes advantage of the first-person perspective inasmuch as the study of consciousness utilizes a subject's "heterophenomenology." A heterophenomenology is constructed by videotaping and soundtaping and electroencephalographing a subject, preparing a transcript from the soundtape, and interpreting the resulting text as a record of speech acts (i.e., treating the subject as a rational agent). The resulting text is the subject's "heterophenomenological world." (Dennett 1991, pp. 74–81) Then, to determine the accuracy of the subjects' reports, the investigator checks the "'defining' properties of the items that populate [the subjects'] heterophenomenological worlds" against the "real goings-on in people's brains." (Dennett 1991, p. 85) Brain events are presumed to be "the real referents of the beliefs we express in our introspective reports." (Dennett 1991, p. 85)³¹

What is of interest here is that the heterophenomenology cannot have broken free of the first-person perspective. For, presumably, the original soundtape from

which the heterophenomenology was prepared contained numerous "I*" sentences by which the subject attributes first-person attitudes to herself*. The first-person perspective is not left behind by a heterophenomenological text that renders Jane's words, "I wish that I* could have seen John once more," for example, as "Jane wishes that she* could have seen John once more." The way to avoid recognizing a first-person perspective would be to render Jane's words as "Jane wishes that Jane could have seen John once more," but, as we have seen repeatedly, a genuinely third-person rendition of Jane's utterance would not be accurate.³² On the other hand, if Dennett were to accept attributions of the first-person perspective to Jane as being consistent with his insistence on the third-person perspective, then he, too, would need a naturalistic account of the first-person perspective.

(C) Finally, there is another angle on self-consciousness—this one associated with Dennett as well. We can think of a very sophisticated kind of self-consciousness as construction of a narrative self. When philosophers speak of a self—as Dennett's self as center of narrative gravity, or Owen Flanagan's self as an emergent set of models—they are considering a much thicker concept than what I mean by "the first-person perspective." (Dennett 1991; Flanagan 1992) Indeed, they tend to skip over the first-person perspective altogether. Flanagan, for example, distinguishes a weak sort of self-consciousness enjoyed by any beings, even infants, that are subjects of experience, from a strong sort of self-consciousness that requires a temporally extended soliloquy or dialogue. We "are self-conscious in a deep way, for example, when we are engaged in figuring out who we are and where we are going with our lives." This sort of self-consciousness involves, Flanagan

says, "thinking about one's model of one's self, or as I shall say for simplicity, the self." (Flanagan 1992, pp. 194–95) Clearly, "thinking about one's model of one's self" must be understood as thinking about one's model of oneself as oneself, and not just as thinking of a model of someone-who-is-in-fact-onself.³³ The point again is that a view of self-consciousness is demonstrably false unless it presupposes the first-person perspective.

The idea of a self is much richer than the idea of a first-person perspective. A self is the locus of personal integrity and coherence, but such a self is not required for a first-person perspective. A. R. Luria wrote of a soldier, Zasetsky, who suffered a brain wound in World War II, and who desperately tried to recover his self, his source of identity. (Luria 1972) What Zasetsky had lost was a coherent and comprehensive story of his life, of which he was the subject; but he had not lost his first-person perspective. Indeed, his quest to find out "who he was" presupposed a first-person perspective. For a first-person perspective is required in order to be in the position of searching for who one* is. A first-person perspective is necessary, but not sufficient, for an idea of a self in the sense that Zasetsky lost.

On each of these views, either a first-person perspective is presupposed without being explicitly discussed or it is genuinely left out. If it is genuinely left out, as I have tried to show, then the view is subject to counterexamples. So, I think that there is no way for an adequate account of self-consciousness to avoid the first-person perspective.

A CHALLENGE TO NATURALISM

Can the first-person perspective be accommodated by naturalism? The answer depends on what is meant by "naturalism."

If naturalism requires only that there be no appeal to anything supernatural or immaterial, then the first-person perspective is naturalistic. A capacity for a first-person perspective seems no less likely to have been produced by evolution than a capacity for speaking a language or a capacity for solving differential equations. Call this sense of naturalism "weak naturalism." There are more robust construals of "naturalism," however, incompatible with the relaxed ontological pluralism that seems plausible to me. A more robust naturalism would be the view that what exists is only what is countenanced by the natural sciences. This more robust ontological naturalism is often allied with various projects of "naturalization." These projects aim to provide full and sufficient explanations of various intentional, semantic, and mental phenomena in nonintentional, nonsemantic and nonmental terms. It is this robust naturalism, in both its ontological and epistemic versions, to which I suggest that the first-person perspective presents a challenge.

At first blush, it may seem that a robust or reductive naturalist would have no difficulty with the first-person perspective. A functionalist, for example, may suppose that someone with a first-person perspective simply has a mental symbol that plays the functional role of "I*." So, if functionalism is otherwise satisfactory,³⁴ the first-person perspective may be thought to present no additional difficulties. Alternatively, the reductive naturalist may recommend invoking some sort of special self-referential concept in higher-order thoughts about oneself. These suggestions simply relabel the problem without solving it. The problem is not with concepts in general, nor with indexical concepts (like "here" and "now") in particular. What is at issue is a particular *ability*—the ability to conceive

of oneself as oneself*, which, as we saw in the case of Oedipus's motivation for his self-blinding, goes beyond the ability to have self-locating beliefs of the kind that Perry discussed. To show how a robust or reductive naturalism could accommodate the first-person perspective, one would have to give a nonintentional and nonsemantic account of the conditions under which an individual has that ability. I am not claiming that this cannot be done, but only that it is a challenge that robust naturalists have not taken up.

In conclusion, the challenge to naturalists is this: Either be satisfied with weak naturalism, or show how the first-person perspective can be accommodated by robust naturalism, or show how to dispense with the first-person perspective. I have given reasons to think that a complete account of human behavior will have to countenance the first-person perspective. So, I think that the options reduce to two: be satisfied with weak naturalism, or show how the first-person perspective can be accommodated by robust naturalism. But I may be wrong; so I await a naturalist's reply.³⁵

University of Massachusetts, Amherst
Received October 1, 1997

NATURALISM

pective be ac-
n? The answer
v "naturalism."

NOTES

1. Some philosophers say that no noncircular definition is possible. For example, see Searle 1992, p. 83, and Block 1995, p. 230.
2. For arguments on each side of this question, see Searle 1992, and Rosenthal 1986.
3. For a discussion of this issue, see Shoemaker 1996.
4. For arguments on each side of this question, see Flanagan 1992, and McGinn 1991.
5. For discussions of this issue, see Levine 1993 and van Gulick, 1993.
6. Several philosophers have taken up and emphasized various aspects of first-person reference, but none has emphasized the importance of Hector-Neri Castaneda's "I*" phenomena. See Nagel 1979; McGinn 1991; Searle 1992.
7. There have been many discussions of subjectivity and related issues in recent years. One of the most prominent is Thomas Nagel's the first-person perspective in terms of something that "it is like to be." See Nagel 1979, and Nagel 1986.
8. This hypothetical remark by the dog is intended to be theory-neutral with respect to what "representations" (if any) may be "tokened" in the dog's brain.
9. For a discussion of (merely) perspectival phenomena, see Perry 1979. David Lewis and Roderick Chisholm are often mentioned together as both giving accounts of all attitudes in terms of self-ascription. I suspect that there may be an important difference between them. Since Lewis treats all indexical attitudes, first-person or not, in the same way, I suspect that for him there is no more to the first-person perspective than perspectivalism (what I'm calling weak first-person phenomena). I think that Chisholm, on the other hand, takes his direct attribution to require what I'm calling strong first-person phenomena. See Chisholm 1981 and Lewis 1979.
10. The *New York Times* (April 22, 1997, C9) reported that recent research by Marc D. Hauser at Harvard University suggests that certain monkeys—e.g., cotton-top tamarins of South America—also can recognize themselves.
11. It would be very interesting to compare brains of chimpanzees and brains of monkeys who apparently cannot learn self-recognition to see whether there are structural differences that can account for the apparent cognitive difference.
12. Wechkin et al. (1964a), (1964b). My description in the text, along with the quotation, comes from Rachels (1976), p. 215.
13. If Gallup's chimpanzees are self-conscious at all, the development of their self-consciousness is not part of their normal socialization. Rather development of self-recognition in chimpanzees requires direct intervention by members of another species. Moreover, the chimpanzees' self-consciousness, if that is what it is, is bound to the circumstances in which it was developed; it is not available for integration into a variety of attitudes independent of the situations that originally provided evidence of their self-recognition. Finally, it is not clear to what extent the chimpanzees' self-recognition is a conceptual ability as opposed merely an ability to discriminate. For these reasons, I do not take chimpanzees with a capacity for self-recognition to have a full-blown first-person perspective.
14. Hector-Neri Castaneda, who did pioneering work on making and attributing first-person reference, used "he*" to mark the reflexive use of "he (himself)." See Castaneda 1966 and Castaneda 1967. For a study of philosophy from a first-person point of view, see Matthews 1992.

15. It does not matter what predicate is substituted for "F." "I think that I am tall" indicates that the speaker can think of herself as herself* just as well as "I think that I will have nightmares."

16. Strawson was right that states of consciousness are ascribed to the very same things as corporeal characteristics; and that thing is the person. Strawson 1963, p. 98.

17. Wittgenstein 1958, pp. 66f. Also, see Shoemaker 1968, and Shoemaker 1986.

18. On this point, I agree with Strawson 1963, viii, who held that a "condition of the ascription of states of consciousness to oneself is ability to ascribe them to others."

19. My point here would not be affected by the truth of innateness hypotheses proposed by cognitive scientists, including Chomsky and Fodor.

20. The assumption that a particular concept may be individuated by inferential role construed solipsistically may rest on a picture like this: a homunculus in my head is watching an internal monitor, where the interpretation of what is on the monitor is determined independently of what the monitor is hooked up to. Such a picture would be totally misleading.

21. Compare: What would count as a solution to an "equation" that consisted of nothing but variables? Or worse: what would count as a certain mark's being a variable, or in virtue of what would the whole string count as being one equation rather than another, or even as being an equation at all?

22. Geach 1957, p. 118. In Geach 1972, Geach considers the indirect reflexive pronoun in "believes that he himself is clever."

23. If we construe Descartes's premise in the cogito, not as "I think," but as "I am certain that I* am thinking," then it does support the conclusion that I exist. But: (1) given his skeptical assumptions about contingent things apart from himself, Descartes is not entitled to either premise. He cannot just help himself to the concept of thinking. (2) Even if I am a thing that thinks (as I agree that I am), I am not an immaterial soul. (3) "I" and "I*" refer to the same thing: me, a person. A person is, in some sense, a substance: an embodied thinking substance with a first-person perspective.

24. Nor does my point concern the semantics of first-person sentences. For example, John Perry would describe the change from third-person to first-person perspective as change of belief state, not change of proposition believed. Steven Boer and William Lycan have argued that there is no semantic difference between first- and third-person perspectives, only a pragmatic difference. See Perry 1979 and Boer and Lycan, 1986.

25. In a footnote, Rosenthal says, "It is not sufficient that the report be about somebody who happens to be oneself. Rather, the report must be about oneself, as such; that is, it must be a report that the being that is in the mental state is oneself." (Rosenthal 1991, p. 476, n. 12) In other footnotes, Rosenthal mentions some of the relevant literature about the difference between believing of someone who turns out to be oneself that he is F, and believing that he* is F. See, e.g., Rosenthal 1993, p. 216 n. 10.

26. Although Rosenthal requires that the associated higher-order thoughts be independent of *conscious* inference, he explicitly allows that "non-conscious inference may well underlie the presence of the higher-order thoughts that make mental states conscious. Such non-conscious inferences are not precluded here, since they would not interfere with the intuitive immediacy of such consciousness." Rosenthal 1993, fn. 18, p. 219

27. Rosenthal, I believe, would agree since he rejects the non-first-person interpretation of his view. See note 25.

example, see Searle

that 1986.

Ginn 1991.

first-person reference,
phenomena. See Nagel

ent years. One of the
something that "it is

with respect to what

Lewis and Roderick
les in terms of self-
Since Lewis treats
him there is no more
first-person phenom-
to require what I'm

Marc D. Hauser at
of South America—

is of monkeys who
differences that can

e quotation, comes

their self-conscious-
self-recognition in
Moreover, the chim-
es in which it was
ndent of the situa-
not clear to what
erely an ability to
self-recognition to

g first-person ref-
66 and Castaneda
as 1992.

28. Compare the KK-thesis in epistemology: If S knows that p, then S knows that she (S? she*?) knows that p. One may suppose that (a) and (b) are related by an analogous principle: If S has the thought that p, then S has the thought that she* has the thought that p. But this could not be Rosenthal's view since he does not think that all conscious states are introspectively conscious. Moreover, the discussion of the first-person perspective shows that the principle analogous to the KK-thesis is false (as is the KK-thesis itself).
29. The need for a reductive account of the first-person perspective is obscured by a tendency to construe the introspective consciousness of a mental state M as one's having the thought that one has the thought that one is in M. Formulated in this way, it appears that the third-order mental state is built up from the first-order mental state by some mechanical means like iteration. Thus, it is not surprising that the need for a reductive account of the first-person perspective has gone unnoticed.
30. To try to accommodate the first-person perspective by hypothesizing that there is a special first-person psychological mode of presentation under which one represents oneself as oneself* is only a relabeling of the distinction from a functionalist point of view. See comments in the final section.
31. Disregard the fact that Dennett's views on consciousness seem quite at odds with his views on intentionality.
32. For what it's worth, I think that heterophenomenology is an excellent way to study consciousness. I just wouldn't claim that it avoided the first-person perspective. Nor would I try to map "items that populate heterophenomenological worlds" onto items in the brain.
33. A number of prominent philosophers take self-consciousness to be access to a self-model. As David Chalmers put it, self-consciousness is "our ability to think about ourselves, our awareness of our existence as individuals and of our distinctness from others. My self-consciousness might be analyzed in terms of my access to a self-model or my possession of a certain sort of representation that is associated in some way with myself." (Chalmers 1996, p. 27) For reasons that I have rehearsed several times, "my access to a self-model" cannot simply be my access to a model of myself from a third-person point of view, but of myself as myself*, from a first-person point of view.
34. I have argued against functionalism on other grounds. See Baker 1987 and Baker 1985.
35. Many thanks are due to Gareth B. Matthews and to Katherine Sonderegger, who read several drafts of this paper and made helpful suggestions.

BIBLIOGRAPHY

- Armstrong, David M. (1968). *A Materialist Theory of the Mind*. London: Routledge and Kegan Paul.
- Baker, Lynne Rudder (1981). "Why Computers Can't Act." *American Philosophical Quarterly* 18: 157-163.
- Baker, Lynne Rudder (1985). "A Farewell to Functionalism." *Philosophical Studies* 48: 1-13.
- Baker, Lynne Rudder (1987). *Saving Belief: A Critique of Physicalism*. Princeton: Princeton University Press.
- Baker, Lynne Rudder (1995). *Explaining Attitudes: A Practical Approach to the Mind*. Cambridge: Cambridge University Press.

at she (S? she*?)
 ipple: If S has the
 his could not be
 tively conscious.
 analogous to the

by a tendency to
 thought that one
 ird-order mental
 e iteration. Thus,
 pective has gone

there is a special
 eself as oneself*
 omments in the

is with his views

at way to study
 Nor would I try
 brain.

a self-model. As
 s, our awareness
 eousness might
 sort of represen-
 or reasons that I
 ecess to a model
 irst-person point

Baker 1985.

who read several

edge and Kegan

hical Quarterly

lies 48: 1-13.

eton: Princeton

the Mind. Cam-

- Block, Ned (1995). "On a Confusion about a Function of Consciousness." *Behavioral and Brain Sciences* 18: 227-247.
- Boer, Steven E. and Lycan, William G. (1986). *Knowing Who*. Cambridge, Mass.: MIT/Bradford.
- Bower, I. G. (1974). *Development in Infants*. San Francisco: Freeman & Co.
- Castaneda, Hector-Neri (1966). "He: A Study in the Logic of Self-Consciousness." *Ratio* 8: 130-157.
- Castaneda, Hector-Neri (1967). "Indicators and Quasi-Indicators." *American Philosophical Quarterly* 4: 85-100.
- Chalmers, David J. (1996). *The Conscious Mind: Toward a Fundamental Theory*. Oxford: Oxford University Press.
- Chisholm, Roderick (1981). *The First Person: An Essay on Reference and Intentionality*. Minneapolis: University of Minnesota Press.
- Dennett, Daniel C. (1991). *Consciousness Explained*. Boston: Little Brown and Company.
- Flanagan, Owen (1992). *Consciousness Reconsidered*. Cambridge, Mass.: MIT/Bradford.
- Frankfurt, Harry (1971). "Freedom of the Will and the Concept of a Person." *The Journal of Philosophy* vol. lxxviii, no. 1: 5-20.
- Gallup, Gordon, Jr. (1977). "Self-Recognition in Primates: A Comparative Approach to the Bidirectional Properties of Consciousness." *American Psychologist* 32: 329-338.
- Geach, Peter (1957). *Mental Acts*. New York: Humanities Press.
- Geach, Peter (1972). "On Beliefs About Oneself." In *Logic Matters*. Oxford: Blackwell, pp. 128-29.
- Kaplan, David (1988). "Demonstratives." In *Themes from Kaplan*, Joseph Almog, John Perry, Howard Wettstein, eds. Oxford: Oxford University Press.
- Levine, Joseph (1993). "On Leaving Out What Its Like." In *Consciousness: Psychological and Philosophical Essays*. Oxford: Basil Blackwell, pp. 121-136.
- Lewis, David (1979). "Attitudes *De Dicto* and *De Se*." *The Philosophical Review* 88: 513-43.
- Lycan, William G. (1987). *Consciousness*. Cambridge, Mass.: MIT/Bradford.
- Lycan, William G. (1996). *Consciousness and Experience*. Cambridge Mass.: MIT/Bradford.
- Luria, A. R. (1972). *The Man With a Shattered World*. New York: Basic Books.
- Matthews, Gareth B. (1985). "The Idea of a Psychological Organism." *Behaviorism* 13: 37-51.
- Matthews, Gareth B. (1992). *Thought's Ego in Augustine and Descartes*. Ithaca: Cornell University Press.
- McGinn, Colin (1991). *The Problem of Consciousness*. Oxford: Basil Blackwell.
- Nagel, Thomas (1979). "What is it Like to Be a Bat?" In *Mortal Questions*. Cambridge: Cambridge University Press, pp. 165-180.
- Nagel, Thomas (1986). *The View from Nowhere*. New York: Oxford University Press.
- Perry, John (1979). "The Problem of the Essential Indexical." *Nous* 13: 3-22.
- Rachels, James (1976). "Do Animals Have a Right to Liberty?" In *Animal Rights and Human Obligations*, Tom Regan and Peter Singer, eds. Englewood Cliffs, N.J.: Prentice-Hall, pp. 205-223.
- Rosenthal, David M. (1993). "Thinking That One Thinks." In *Consciousness: Psychological and Philosophical Essays*, Martin Davies and Glyn W. Humphreys, eds. Oxford: Blackwell, p. 199.
- Rosenthal, David M. (1991). "Two Concepts of Consciousness." In *The Nature of Mind*, David M. Rosenthal, ed. Oxford: Oxford University Press, pp. 462-477. [Reprinted from *Philosophical Studies* 94 (1986): 329-59.]
- Russell, Bertrand (1945). *A History of Western Philosophy*. New York: Simon & Schuster.
- Searle, John R. (1992). *The Rediscovery of the Mind*. Cambridge, Mass.: MIT/Bradford.
- Shoemaker, Sydney (1996). "Self-Knowledge and the 'Inner Sense'." In *The First-Person Perspective and Other Essays*. Cambridge: Cambridge University Press, pp. 201-268.
- Strawson, P. F. (1963). *Individuals: An Essay in Descriptive Metaphysics*. Garden City, N.Y.: Doubleday Anchor Books.
- Sartre, Jean-Paul (1966). *Being and Nothingness*. New York: Washington Square Press.

- Shoemaker, Sydney (1968). "Self-Reference and Self-Awareness." *The Journal of Philosophy* 65: 555-67.
- Shoemaker, Sydney (1986). "Introspection and the Self." In *Studies in the Philosophy of Mind*, Midwest Studies in Philosophy, Vol. X, ed. Peter A. French, Theodore E. Uehling, Jr., Howard K. Wettstein. Minneapolis: University of Minnesota Press, pp. 101-120.
- Stern, Daniel N. (1985). *The Interpersonal World of the Infant*. New York: Basic Books, Inc.
- van Gulick, Robert (1993). "Understanding the Phenomenal Mind." In *Consciousness: Psychological and Philosophical Essays*. Oxford: Basil Blackwell, pp. 137-154.
- Wechkin, Stanley; Masserman, Jules H.; and Terris, William Jr. (1964a). "Shock to a Conspecific as an Aversive Stimulus." *Psychonomic Science*, 1: 47-48.
- Wechkin, Stanley; Masserman, Jules H.; and Terris, William Jr. (1964b). "'Altruistic' Behavior in Rhesus Monkeys." *The Journal of Psychiatry*, 121: 584-85.
- Wittgenstein, Ludwig (1958). *The Blue and Brown Books*. Oxford: Oxford University Press.