

How to Think About Thinking: A Preliminary Map

J. Michael Russell

California State University, Fullerton

I discuss three orientations to what the word "thinks" might mean, the mentalist, the behavioral, and the avowal orientations, and explain some variations of each. I urge that mapping out thinking in this way allows us to examine some important issues that escape us with more familiar theories (e.g., mind-brain identity theory, the thesis of intentionality, etc.) and that these have important implications for theorists in the social sciences. I argue that psychological behaviorists often turn out to be philosophical mentalists in disguise, and that a position of philosophical behaviorism is profoundly different from the sort of behaviorism familiar to psychologists.

In what follows I shall try to map out some ways to think about what *thinking* is. By in large, what I have to say about thinking will not be new to persons familiar with philosophical analysis in the last thirty years, and I would like it to be comprehensible to persons without such backgrounds. This will require that I consider quite complex positions in a most abbreviated form. I draw a map of territory covered by others, with the objective of orienting newcomers to this topic, and of posting reminders for how to find one's way about in an area where it is easy to become lost. I would not want to give the impression that I am somehow presenting in summary form a consensus which recent philosophers have achieved on this topic; there is no such consensus. Rather, I present a network of themes and variations which I believe should be interesting and challenging to theorists in other fields. I will want to show (in a paper yet to be written) how this map may be of use to possible theories about the nature of psychotherapy, and will offer some hints about this in the present text.

I shall concentrate on variations of three orientations to what thinking is, and to what words like "thinking," "thinks," and "thought" *mean*. Each has its merits and its weaknesses. One treats thinking in terms of private mental events, inner processes or occurrences or states. According to a second orientation, what it is to think such-and-such is to have a tendency to behave in a certain range of ways appropriate to the application of that term. And according to a third orientation — a close cousin of the second — the person who thinks such-and-such is the per-

son who is prepared to *say* that he or she thinks such-and-such.

Overview

The orientations to be considered may be dubbed, respectively, the *mentalist*, the *behavioral*, and the *avowal* orientations. I speak of "orientations" so as to indicate that I am speaking loosely of various views which I regard as related; specific theorists may not fall very neatly under these rough headings. However, among traditional philosophers, Descartes and Locke stand out as mentalists; probably most theorists and laypersons belong here. Wittgenstein (1958 a, b) and Ryle (1949) are prominent figures among a number of analytic philosophers who favor combinations of the behavioral and avowal orientations. (It is extremely important to note — for readers trained in the social sciences — that the behavioral orientation of many philosophers differs profoundly from what psychologists associate with the word "behaviorism.") In my opinion, various existentialist, e.g., Sartre, fall more under the behavioral than the mentalist heading; but this claim requires a defense I shall not pursue here.

Each of these orientations prescribes an approach to what it *means* to say that a person thinks such-and-such. They seek to identify the criteria — and not merely the symptoms — for what the word "thinks" means. For example, when we say of a particular man that he thinks that it is raining outside, what does this expression mean? According to the mentalist, the expression applies if (and only if) this man has some relevant sort of process, state, or occurrence of something called (the relevant) "thought," "in his mind." But the behaviorally oriented theorist, perhaps motivated by seeming failures in the mentalist's position, will attempt to construe the meaning of "that man thinks it is raining outside" in terms of the man's behavior, and the variations in behavior which presumably would be forthcoming depending on variations in circumstances. Thus the person who thinks that it is raining will be — by virtue of the (supposed) meaning of the expression — a person disposed to doing such things as picking up an umbrella and overcoat when going out of doors.

These dispositionally oriented analyses turn out to be indefinitely complex "if . . . then" conditional statements. *If* the man isn't just returning the umbrella to a friend, and *if* he doesn't carry the umbrella in the manner of someone oblivious to the rain, and if . . . etc. (there may be an indefinitely long list of "ifs"), then he thinks that it is raining. Or suppose he does not open the umbrella when he is getting wet in the rain: then *if* we have reason to suppose he wants his oversized pants to get soaked and shrink (presumably this "he wants . . ." would also be analyzed in terms of relevant behavioral dispositions), etc., then he's a man who thinks it is raining. The idea is that these indefinitely complex descriptions of what the man *would* do under various contingencies pro-

vide the very meaning of the expression "he thinks it is raining".

Each of these orientations trades on the weaknesses of other positions. It is a weakness of the behavioral position, it seems, that no matter what the man does it will be possible for an imaginative person to construe it in more than one way. Even a very complicated description of how someone packs an umbrella after looking out the window at the sky, etc., may seem to add up to thinking that it is raining when some alternative account of what they "really" think — even a bizzare or far-fetched alternative — is available. Maybe the man "really" thinks that the umbrella will serve to fend off the devils whose little feet (he thinks) can be heard pattering along the sidewalk!

Since there can be such disputes about how to interpret a person's behavior, the avowal oriented theorist siezes upon that which settles disputes, and holds that it is this which provides the key to the meaning of attributing a specific thought to a person. What the man himself says he thinks is in fact what typically settles disputes. So, according to this avowal orientation, the person who thinks it is raining is one who, in addition to other relevant behavioral dispositions, will say upon being questioned (and supposing there is no deceit or error), "I think that it is raining." Having the disposition to make such a (sincere) avowal, plus other relevant behavioral dispositions, is supposed to provide the very meaning of "thinks that it is raining." ("Sincerity" presumably must itself, in turn, be analyzed into relevant behavioral dispositions.) Meanwhile, the mentalist will claim it is obvious to "common sense" that it is what goes on in the man's mind that makes all the difference.

Connections with Other Theories

Before I begin to develop (and juggle) aspects of these orientations, I want to urge that considering them will help us to focus on some questions which other kinds of theories often fail to bring out. For instance, they help us to see certain options in how we may think about thinking which are not brought out by "identity" theories. Theorists who argue for the prospect of identifying thinking (and consciousness) with some brain process are often unclear about just *what it is* that they are prepared to identify with a brain process. For instance, if we were persuaded that thinking it is raining outside may turn out to be identical with some complicated neurophysiological happening in the brain, it would still have to be made clear *what* we were identifying with this brain process. Even pretending that we are clear enough about what is meant by "brain process" and "identical with," it is *not* clear enough just *what* is supposed to be identical with this brain process. Is it a "mental occurrence"? Is it a complex behavioral disposition? Is it a disposition to make certain linguistic avowals? *Prima facie*, a mind-brain identity theory would be compatible with any of these approaches to what thinking is. But we can't get far trying to identify brain process with the condi-

tion of a person who *thinks* such-and-such until we make an independent decision about what we mean by saying a person thinks such-and-such.

There is a similar problem with the famous "thesis of intentionality." According to this theory, it is essential and definitive of all mental states, in contrast to all physical states, that they are "of" or "about" or "directed toward" something. Supposedly, to have a thought or belief or wish or intention is always to have a thought (etc.) *about* something. The essential nature of the mental is that it has this feature of "aboutness." Milk bottles and stones are not *about* something, whereas thoughts are always thoughts about this or that. Yet like the identity theory, this thesis of intentionality fails to tell us *what it is* which supposedly has this feature of aboutness or intentionality. Is it the mental *event* cited by the mentalist? Is it something inherent in the behavioral disposition? Is it a fact built into the language of avowal?

Or again, consider the *epiphenomenalist's* theory. The epiphenomenalist believes that our mental expressions (hence words like "thinks") refer to something which cannot be adequately reduced to a description of the "merely physical." So they admit that when we say what a man thinks we discuss something different from what is going on with his body. *But*, the epiphenomenalist contends, this mental dimension is the mere irrelevant by-product of the purely physical world. What we think is entirely produced by the causally determined physical world, and yet in turn exercises no real influence on that world. What we think is a mere by-product, and like the steam escaping an engine, it "turns no wheels." And yet for this theory there is once again the question: *What* is it which is supposedly this mere by-product? Is it the events cited by the mentalist? Is it our behavioral dispositions? Is it what we would *say* about ourselves?

These last paragraphs should bring out the importance of the orientations I have sketched, for they raise questions which cut across more familiar theories about thinking.

Variations on the Mentalist Orientation

Theorists of this persuasion are liable to account for what it is for a man to think that it is raining in terms of the "inner occurrence" of something to be called "the thought that it is raining." What sort of occurrence? Traditional mentalists like Descartes will maintain that we are all familiar in a "direct" way with the existence and nature of our thoughts. But can anything further be said about the nature of these? I can think of two main routes to try here. According to one, mental occurrences are inherently unlike and irreducible to anything resembling publicly witnessable events. For reasons which will become evident, I will call this the *non-analogue* view. Evidently it will be difficult to say much about such mental events; one must resort to what are metaphors of an admittedly crude and misleading sort in an effort to refer to

something supposedly plain to each of us in our own "private" experience. Many people seem to believe that this is the right sort of way to discuss the nature of thinking; we are not to expect too much.

According to another route, mental occurrences are what I shall call *phenomenologically analogous* to publicly witnessable occurrences. Silently speaking on some matter is a phenomenological analogue of talking about it aloud. Visually imaging something is a phenomenological analogue of seeing it happen in the public world. The mentalist may believe that thinking it is raining consists in having relevant occurrences on the order of silent speech or mental imagery. In contrast, the non-analogue theorist may contend that such events as silent speech or visual imagery are outgrowths of what it is to think that it is raining, and believe that thinking this is a kind of occurrence, but not believe that thinking is a kind of occurrence which resembles any familiar public events.

What is sometimes used as an argument for the behaviorist orientation may actually only indicate that something is wrong with the orientation of the mentalist who wants to represent "thinking that such-and-such" as consisting of events like mental imagery which are phenomenologically analogous to events in the public world. It is argued that thinking that it is raining cannot consist in having a mental image of rain. For a person might have such an image, yet the image not be "of" rain but "of" something which "looked" similar, such as water from a hose, or slow moving drops of glycerin from a chemical fountain. Or it might be that the person indeed had the image of rain, but that this was part of an idle day-dream and did not represent his belief (thought) that it was raining outside. (Thinking *of* Aunt Tilly chewing bubble gum differs from thinking *that* Aunt Tilly is chewing bubble gum.) It is a fact that the having of a mental occurrence which is the analogue of some public episode cannot be identified with what it is to think that such-and-such is the case, for the reason that such events can be construed variously in different circumstances. But someone who subscribed to a non-analogue mentalism might maintain that while images and words are ambiguous, what it is to *think* something is not ambiguous.

Inner States

In my sketch of mentalism I have had to use the term "inner." The term is used metaphorically, for it is not meant that thinking is something which takes place in some locatable cavern. I believe the point of using the word "inner" is to stress the idea that what a person thinks is accessible only to that person in an "immediate" way; to call it "inner" is to contrast it with publicly observable behavior, with which it is only contingently connected. Others have to take a person's word for it on what he or she thinks, or else infer its existence "indirectly;" a person knows what he or she thinks, while others can only guess.

In discussing the mentalist position I have been running together the terms "mental occurrence" and "mental state." I must confess to being unsure of what a mental *state* is supposed to be. Perhaps it is a prolonged occurrence. So viewed, it would be something ongoing, which might have begun sometime, and might end sometime. In this sense a body of water might be said to be in a frozen state. If a person's thinking that it is raining is said to be a state, it is portrayed as a "something" which is "there" ("in the mind") and such as will be "found" by an "introspective look". It compares with how I shall find ice cubes in the refrigerator whenever I happen to look.

A different rendering of the term "state" would come from treating the idea of a mental state as the *disposition* for certain privately accessible events called "thoughts" or called "having the thought that it is raining" to take place (occur). On this view a woman who truly thought it was raining would not necessarily have at a given moment an ongoing event-like business of thinking that it was raining. Perhaps the woman is having a nap — a dreamless nap — in the belief that as it is raining there is nothing interesting to be done with the afternoon. But when appropriate (and when awake, presumably) she would have whatever sort of occurrence (or occurrences) which constitutes what it is to be thinking that it is raining, and this would be of the sort that the mentalist claims.

A mentalist who gave dispositional analyses for "thinks" might subscribe to either a public analogue view or a non-analogue view of the nature of the occurrences or states in which thinking supposedly consists. In either case, it seems not to have always been recognized that a mentalist could employ the concept of a disposition. One might have assumed (as I have tended to do in the past) that this term was a tool only of use to the behavioral theorist. However, maybe the concept of a disposition isn't as interesting in a mentalist's orientation as it is in a behavioral orientation. For when the concept of a mental state is understood in terms of there being a disposition for certain mental occurrences, the really interesting constituents of these inner states are the inner occurrences, and what we would like to have is a better portrayal of them.

Objections to Mentalism

Why would anyone bother to reject mentalism? I have said that I wanted this paper to be comprehensible to persons unfamiliar with recent philosophical analysis, and such persons well may not see why they should struggle with views which go so much against what seems so plain to "common sense." Much of the intellectual world is not yet even aware of the enormous amount of critical philosophical discussion of this idea that each of us knows the nature of thinking from an immediate access to our own inner lives. Yet contemporary criticism of this idea may have the most serious implications of any direction in recent philosophy. Here I must confine myself to a very summary mention of some reasons for

questioning mentalism.

Here first is a phenomenological objection to mentalism. Theorists of this orientation tend to maintain (perhaps unnecessarily) that our thoughts exist within us in such a way that we are each aware of the presence of our thoughts. Yet people who answer affirmatively to the question "Do you think it is raining?" will not always allow that when they answer such questions they do so by making reference to an inner state or occurrence of which they are or were aware. It just does not seem to be a simple fact of experience that we are aware of thoughts within us which enable us to say what we think.

Secondly, mentalists tend to hold (again, perhaps unnecessarily) that our capacity to deal with the world intelligently is made possible by our actions being preceded by the occurrence or existence of our thoughts. Now as was just noted, it is less than patent that experience bears this claim out. But anyway, as Gilbert Ryle (1949) has argued, such a claim tends to generate an absurd "infinite regress." For consider the claim that before you could act capably you must think about what you are doing. Thinking about what you are doing is itself an activity which may or may not be done well. So if it were true that before you could do something well you had to think about it, it would also be true that before you could think about it (or think about it well, i.e., carefully, capably) you would have to think about thinking about it. And, in turn, you would have to think about thinking about thinking about it. As there is no end to this regress, there must be something absurd about the assumption that generates it, namely, the idea that before doing anything well you must think about doing it.

The preceding argument works against the notion that thinking is something which precedes each and every act; a view generated by the grammatical fact that one of the expressions we use for indicating that we did something well is to say we thought about what we were doing. Here is a more general criticism of mentalism. The mentalist maintains that the *meaning* of "thinks" is understood by each of us by reference to our own thoughts, and that these are occurrences or states to which the person whose thoughts these are has sole access. But if each of us knew only from our own case what it is to think something, then we would have no good reason to believe in, and indeed not even means to comprehend the very idea of another person's thinking. The point here is that we obviously do know what other people mean when they say they think that it is raining, and yet we could not know this if the mentalist's thesis about the meaning of "thinks" were correct.

It is often supposed that since other people have behavior analogous to our own, we are entitled to infer in them the existence of mental states or processes comparable to what supposedly we each find within ourselves. But this argument from analogy is untenable. Just suppose that there were something about mailboxes which, on principle, made it utterly impossible for anyone to ever inspect the contents of any mailbox but

their own. And suppose, one morning, that I open my mailbox and find something in there which I call an "Easter egg." I dance with excitement! Then I notice my neighbor dancing excitedly upon having inspected his mailbox. Have I any reason whatsoever for inferring, by virtue of our analogous behaviors, that he has something in his mailbox which is similar to what I have found in my own? Suppose I see similar behavior in everyone on my block, and even suppose that each of them say that they have found Easter eggs in their mailboxes. What is to prevent their mailboxes from being quite different from my own, and their having frogs or marbles or nothing at all? Clearly an inference from one case (my own) to all cases yields a conclusion which is less than certain; indeed, from a methodological standpoint, it would not even seem warranted to say it is "probable" that others have similar contents in their mailboxes. And if we are talking of "contents" in these mailboxes which have the peculiar property that no one but their owners can *ever* inspect them, then the conclusion of this argument from analogy is worse than improbable — it is meaningless.

The argument against the possibility of a "private language" given by Wittgenstein (1958b, paragraphs 241-281) provides a specific and apparently devastating rebuttal to this idea that the meaning of "thinks" could be provided by our individual familiarity with what goes on within us when we think. Here is an abbreviated version of this argument: The use of words in our language is taught. In order to teach the use of words to a child, it must be possible to distinguish between the cases where the child employs the term correctly and where incorrectly. It would be impossible to distinguish correct from incorrect use if using a term correctly were a matter of linking the term with something to which only the instructed and never the instructor had access. So, the use of the word "think" is taught, and this requires instruction and correction. Such instruction would be impossible if the word were to get its meaning from referring to what the instructor could never have access to "within" the instructed. And so the word "thinks" must not refer in its meaning to this inaccessible something!

Varities of Behaviorism

This is not the place to discuss the (above) private language argument in any sort of detail; it is sufficient if presenting it provides reason for questioning the mentalist orientation (and the vast amount of popular opinion and psychological theory which in one or another form endorses a mentalist orientation) and looking for alternatives. The behavioral orientation appears to provide a viable alternative, since if the meaning of "thinks it is raining" consists in a complex behavioral disposition then it appears possible to instruct a person on using the language. The teacher can indicate the sorts of behaviors to which the term "thinks" applies, and can check on whether the child applies the term appropriately.

What I have called the behavioral orientation need not be the same as specific doctrines of behaviorism. It helps to distinguish *methodological* from *metaphysical behaviorism*. The methodological behaviorist believes that for the practical purposes of doing empirical psychology scientists should confine themselves to the variables to which they have access and over which they can exercise control, these being overt behaviors. Methodological behaviorists tend to be unclear about whether they are mentalists who reject the importance of the inner mental life (we might call such persons "pragmatic epiphenomenalists") or whether they reject the very idea that there *is* the sort of mental life the mentalist wants to talk about. The latter are metaphysical behaviorists. (In psychology many theorists who identify themselves as behaviorists are really mentalists in disguise — latent or closet mentalists — because in refusing to incorporate words like "thinks" in their work they tacitly assume that such words do indeed refer to inaccessible private states.) The metaphysical behaviorist addresses himself or herself to the question of the *meaning* of terms like "thinks" and holds to some form of the view that the meaning of this term is a matter of the behavior which "counts" for a case of thinking (such-and-such). It is unfortunate that methodological behaviorists are often blind to the theoretical avenues available to them if only they become clear about where they want to stand on the issue of metaphysical behaviorism. Some methodological behaviorists are mentalist-enough that they believe that such terms as "thinks" ("wishes," "hopes," etc.) refer to private mental states, and then metaphysical-behaviorist-enough to dismiss such notions as having no significance. In this way it happens that behaviorally oriented psychologists often have supposed that they must exclude from their work considerations of what their subjects think (etc.), on the uncritically assumed grounds that such mentalistic words refer to something hidden. An available option would be to construe the meaning of such terms via behavioral dispositions; then, in principle, one might be able to include quite varied mentalistic terminology within one's work and thereby utilize a richer and more significant language. Maybe this utilization of everyday mentalistic language is not feasible for the practical needs of the psychological researcher; maybe the task of sorting out behavior disposition criterion is just too cumbersome. But minimally, it would behoove us to get clear about the reasons for the position we take on this topic.

It may also prove valuable to distinguish *hard-nosed* from *soft-nosed* metaphysical-behavioral orientations. The hard-nosed variety will construe the meaning of "she thinks it is raining" as the disposition to engage in standard sorts of publicly observable ways which purportedly "count for" thinking that it is raining. The soft-nosed behaviorist might be willing to include such non-public occurrences as silent speech or mental imagery, but add that our capacity to understand such occurrences as these, which happen not to be publicly observable, is parasitic on our

ability to understand public occurrences of which these are analogues. I know of no behaviorally oriented philosopher who is not soft nosed in this sense. This seems to have been Wittgenstein's view, since in the *Philosophical Investigations* he says that he does not deny that there are inner processes, but then says, "An 'inner process' stands in need of outward criteria" (1958b, paragraph 580). He also says, "Imagine people who could only think aloud. (As there are people who can only read aloud)" (Paragraph 331). I take this to mean that there are thought occurrences which happen to be private but which could *in principle* be public. The sense of "private" at issue in the "private language argument" is that which is inherently and essentially private; the argument works against identifying the meaning of a term with what is not in principle public, since criterion for identifying whether a word is working rightly can only be taught with respect to something which is in principle publicly identifiable.

Then a metaphysical behaviorist may believe that a word like "thinks" may range over occurrences such as silent speech or mental imagery, but will maintain that the meaning of such terms will be in some important way rooted in their application to public life.

What then, is the difference between a soft-nosed behaviorist who believes in mental occurrences like silent speech, and a mentalist who believes that the mental occurrence of thinking something is phenomenologically analogous to saying it aloud? I suppose it is that the one believes that a term which gets its meaning from its public application comes also to single out events which happen to be (contingently) private, while the other believes that a term which refers to what one initially knows from one's own case comes to apply to behavior one observes in others.

Avowal Orientations

Human speech is, of course, a part of human behavior; so a view which stresses the importance of what a person would avow (say of oneself) will not be distinguishable in a very neat way from other forms of behavioral orientations. However, the reasons for considering avowals separately may be clearer when we recall a reason for stressing behavior (as opposed to private mental occurrences) in the first place. It was argued that if we identify the meaning of "thinks" with privately having the occurrence of a thought, it will always be impossible to know whether to say what another person thinks. But we do sometimes say what others think on the grounds of how they behave. And so — for the metaphysical behaviorist — it becomes tempting to identify the very meaning of "thinks" with the behavioral grounds upon which we base our attributions of this term. The move here consists of identifying the meaning of the term with the criteria for employing it. Now since it seems to be the case that our criteria for settling disputes about what a person

thinks in cases where that person's nonverbal behavior is ambiguous is usually a matter of getting that person to *tell us* what he or she thinks, we may wish to hold that such avowals provide the critical grounds for what thoughts we should attribute to someone.

Since an avowal oriented theory would trade on the important status we actually give to what a person says about what he or she thinks, a means for distinguishing *types* of avowal orientations will be to distinguish different kinds of reasons for regarding avowals as so important.

The most "natural"-sounding reason for giving a person's avowal such important status would be, I suppose, that these avowals are reports of inner mental states or occurrences, and that the person knows best which such occurrences take place within him or her. But of course this isn't an avowal theory at all; it is a type of mentalist theory according to which the meaning of "he thinks such-and-such" consists in having the appropriate such-and-such-ish thought occurrence or state. The difference is a matter of what the theorists takes the word "thinks" to mean; but this can also be brought out by distinguishing "symptom" from "criterion." The mentalist is holding that what people avow is characteristically symptomatic of what thoughts they have within. But an avowal theorist would be maintaining that, given the appropriate surrounding circumstances, the person's avowal is not a symptom or indirect clue to what the person thinks, but *just is* what the person thinks, hence the criterion for attribution.

A second approach to why avowals are so important might be this: While it is true that "thinks" refers to behavioral dispositions, each of us is better acquainted with his or her own behavioral dispositions and history than anybody else, and is therefore in a better position to attribute self descriptions about how he or she is disposed to behave with the language of "I think." "I know what I think because I spend more time with me than anybody else does." But again, this is not a *bona fide* avowal oriented theory. It is an endorsement of a behavioral orientation. It does not maintain that the meaning of "she thinks" is a matter of her being prepared to say "I think," but that it is a matter of how she is disposed to behave, about which she has the best knowledge, but knowledge of the same sort as anyone else might have of her.

Notice that a soft-nosed behaviorist, who viewed some kinds of occurrences such as silent speech to be "contingently private behavior" which could in principle be public, might subscribe to the following: When a man says what he thinks, he remarks on how he is disposed to behave. Part of the behavioral tendencies of his which he knows about and can draw from happen to be private. They could be public, but they are not, since he has, for example, said much only in "silent speech" which he might have said aloud. (If he had said these things aloud this would provide behavioral information which everyone could draw from in characterizing his dispositions.) So when this person says what he thinks

he draws on more information about his observable activities than others have about him, and also draws on his knowledge of "behavior" which, it just so happens, only he has access to.

Many philosophers would reject as absurd the suggestion that when we say what we ourselves think we do so on the basis of *grounds*. We don't have grounds or evidence for saying what we think! We just say it! But I myself do not find this absurd. It seems to me that there are some uses of "think" which serve to characterize or describe behavioral tendencies, and that we learn to apply these uses to ourselves. Of course we do not know about ourselves quite on the same sorts of observational grounds others have. I may be aware that I fidget, but you see me do it, and the difference matters. Yet from the fact that you and I *come in different ways* to know how I tend to behave, it does not follow that we are not both speaking of how I tend to behave when we speak of what I think.

A better representative of an avowal orientation would be a theory which portrayed avowals as *state-expressive*. In contrast to the mentalist orientation according to which an avowal *reports* the existence of an (inner) state or occurrence, this approach would represent what we say, and where we say "I think," as in some way *a part of* what it is to think that thing. There may be differences about what sort of "way" or "part" this is. One possibility would be for the theorist to hold that the avowal is a part of the general package of behavioral tendencies, deserving a special place because of being the least ambiguous part.

Another possible approach would be to treat the avowal as the "natural expression" of a state which might have non-public aspects. For instance, perhaps there is a non-observable component to what it is to think something, and perhaps it is not even in principle public and has aspects which are phenomenologically non-analogous to public behavior; *but*, public behavior, especially speech behavior, would "naturally tend" to be displayed in certain characteristic ways by persons who were in the state that included that component. This approach resembles but differs from the mentalist view that "I think . . ." reports the (inner) presence of a certain thought. For the theory here considered need not maintain that "I think" *means* that there is some inner state, nor partially inner state, nor that it is an attempt to report such a state. The idea, rather, is that the words are a *part* of the state, and "naturally express" it. Similarly: when I am happy I tend to smile. My smiling is not the same as my being happy, but is the "natural expression" of it. My smile is not a *report* about my being happy, but is *expressive* of happiness. At least so it is in the straight-forward case of self-expression. Sometimes I may more make a point of smiling, and this may be fakery or it may be a legitimate exaggeration of the state I am in. Then smiling is more patently a form of communication, but still not a report of an inner state; it is more a matter of taking on the role of one who is unqualifiedly happy. By that analogy, to say what I think is not to report on what is going on inside, but to enter fully (and spontaneously) the role of one

who thinks such things — that being a public role.

In this last perspective we bring more into focus the fact that linguistic avowals must be understood not merely as “natural expressions” but complex practices governed by social conventions and rules. Of course much human behavior is convention-governed, but this is so conspicuously and importantly true of linguistic behavior that it provides a further reason for distinguishing avowal orientations from other behavioral orientations. The thesis which emerges can be put this way: The *meaning* of “he thinks x” derives from the interconnections between this linguistic maneuver and the whole complex network of rules and conventions of what Wittgenstein (1958b) called the “form of life” in which it is immersed. Partly for this reason, the meaning of this expression cannot be reduced either to the private referent claimed by the mentalist, nor to the indefinite list of behavioral dispositions advocated by the behavioral orientation. The meaning of “he thinks x” rests on a certain sort of paradigm case in which he would, under appropriate circumstances, say “I think x.” The conditions for the paradigm must include the man’s being a full-fledged member of the linguistic community, set in a context free of the sorts of difficulties which would be posed by the need for either sham or exaggeration. More problematic cases get their meaning by their parasitic relationship to the paradigms.

In another article (Russell, 1973) I have developed a variant of the avowal orientation in which I compared avowals to what J. L. Austin (1962) called “performative speech acts.” Performatives have the quality that, because of surrounding linguistic conventions, one can make something so just by saying it. I make it so that I promise when I say the words “I promise,” and similarly make a bet with “I bet,” get married with “I do” (in a marriage ceremony), etc. Here it is a matter of the conventions within the language for using a certain expression which explains why persons can only say “I bet” (etc.) for themselves. If saying “I think” is something like *taking a certain stand* via linguistic conventions, that would explain the importance of a person’s avowal (without supposing special access to private states); for a person can take a stand only for himself or herself. “I think” cannot be a pure performative since, according to Austin’s formulation of that notion, a performative utterance is to have no truth value. But because parallels between “I think” and “I promise” are striking, I have urged considering “I think” a “quasi-performative.”

I want to recommend just one more subdivision under the heading of avowal orientations, and it turns on a pretty subtle nuance. Earlier I urged considering “he thinks” as meaning “he would say ‘I think,’” and presented “I think” as an avowal which is *expressive* of what it is to think such-and-such. Now the term “expressive” carries the suggestion that the words of avowal somehow grow out of or are a product of a prior condition or state. So viewed, the avowal “I think” would be the most conspicuous and significant feature emerging from a different and

prior state of thinking. If we go too far in driving a wedge between that "state" and the avowal we lose the whole point of the avowal orientation, which was to identify the meaning of "think" with the avowal. Even short of that, it may help to distinguish between an expressive-avowal orientation which treats the avowal as an outgrowth, and what I would call a *constitutive* analysis which treats the avowal as a (particularly important) *part* of what it is to think such-and-such. So viewed, "I think it is raining" is a constituent (or part) which makes up what it is to think that it is raining, and this differs from the "expressive orientation" (slightly, anyway) because it doesn't carry the hint that the words of avowal somehow emerge from a prior and distinguishable condition. Here my use of "constituent" is indebted to an excellent article by Alexander Sesonske (1968), and I have developed it elsewhere (Russell, 1978a) as a tool for distinguishing between avowals which are constituent of character traits and avowals which are constitutive of affect (feeling).

The Map in Outline: What "He Thinks X" Might Mean

I. *Mentalist Orientation*

"He thinks x" means he has some sort of inner states or occurrences of a relevant sort, to which he has special access.

- (a) *Non-analogue version.* These mental contents are unlike anything in the public world, and are not reducible to, e.g., visual or auditory imagery.
 1. These are occurrences or episodes.
 2. These are ongoing states.
 3. These are dispositions to having (relevant) occurrences or ongoing states.
- (b) *Phenomenological analogue version.* These mental events resemble public events; visualizing something is like seeing it, silent speech is like speaking aloud.
 1. Thinking x consists of having these events.
 2. Thinking x consists of ongoing states from which these events emerge.
 3. Thinking x consists of dispositions for having such events or ongoing states.

II. *Behavioral Orientation*

The meaning of "he thinks x" is to be analyzed into the disposition to display an indefinite list of relevant sorts of behavior under relevant circumstances.

- (a) *Methodological version.* For the practical purposes of science, expressions like "he thinks x" are either not to be used at all, or else analyzed into behavioral dispositions.
 1. This is because such expressions refer to inaccessible events. (This is really disguised or latent mentalism.)
 2. This is because the behavioral dispositions which might

unpack the meaning of "he thinks x" are just too cumbersome for the scientist to work with.

- (b) *Metaphysical behaviorism.* The meaning of "he thinks x" is to be analyzed into the disposition to display . . . etc . . .
1. Hard-nosed variety. The analysis of "he thinks x" is to employ reference to only those behaviors which are straight-forwardly publicly observable.
 2. Soft-nosed variety. Certain events which are not straight-forwardly public behavior, e.g., silent speech, are included in the list of "behavioral" dispositions which unpack the meaning of "he thinks x," on the grounds that these could, in principle, be replaced by straight-forwardly public behavior.

III *Avowal Orientation*

For reasons other than those held by the mentalist and the behaviorally orientated theorist, the leading criterion for saying of a man "he thinks x" should be that he is prepared to avow "I think x."

(a) *Illegitimate versions of avowal orientations.*

1. The reason why the avowal "I think x" is so important is that each person knows best the contents of his or her own mind. This is really a mentalist view.
2. The reason why the avowal "I think x" is so important is that each person is more familiar with his or her own behavior than anyone else. This is really simply the behavioral orientation.

(b) *Versions of avowal orientations.*

1. The avowal is part of the general package of relevant behavior, deserving a special place because it is the least ambiguous part.
2. The avowal is important because of its inter-connections with a whole "form of life" involving complex linguistic rules and conventions. "He thinks x" gets its meaning from paradigm cases where a person would avow "I think x."
 - a. Avowing "I think x" is to be analyzed as a quasi-performative speech act.
 - b. Avowals are constitutive (and not just expressive) of the states they avow.

Preferences and Conclusions

In this paper I have indicated several objections to the mentalist orientation to the meaning of "thinks." In my opinion, these objections should teach us that this orientation does not belong — so to speak — in the center of our map. I believe that mentalism has been vastly over-

rated, but I do not conclude that it has no place on the map at all. To see that it has been over-rated, to challenge the supposition that when we talk about what a person thinks we must thereby commit ourselves to speculations about inherently private states or occurrences, can be tremendously liberating for the psychological theorist. For, it seems to me, there are many theorists in psychology who would *like* to discuss what people think but who refrain from doing so on the grounds that such terminology automatically entails mentalism. We are in a position to challenge that assumption, and to consider the merits of alternative perspectives. I do not believe that "he thinks x" *always* implies that he has (relevant) states or occurrences which are inherently accessible only to him. But I believe that we would go too far in the other direction if we supposed that "thinks" is always and simply a matter of behavioral and/or avowal dispositions set in a context of linguistic practices. I see no reason for supposing that there must be one and only one sort of use for the word "thinks," and only one sort of ground rule governing its meaning. And so my sympathies for the second and third orientations in this paper do not lead me to conclude that it is *never* the case that when we wonder what a person thinks we are wondering what may be going on "inside" that person. Only, I should say, this is not the heart of the concept; this does not belong in the center of the map.

Historically we have assumed that mentalism was the heart of the concept of thinking, and that what a person did and said were to be regarded as peripheral phenomena. I believe it is the other way around. The language games with which we attribute thoughts have their heart and home in what we witness people do. What understanding we may have of what is either contingently or inherently private has a derivative and secondary status.

To take one example of how this applies to a psychological theory, Freud seems to have assumed a mentalistic orientation, and so he seems to have assumed that what a patient avowed was a *clue* about that patient's privately accessible conscious thoughts, while behavior and verbal slips and free associations were clues about inner states or occurrences which were equally private, in principle accessible directly only to the patient, and in fact inaccessible to the patient as well. Freud would have acknowledged that he was speculating when he talked about what his patients unconsciously thought, but he supposed that he was speculating about what was going on inside them. I believe he was speculating about how his patient's deeds and historical (but not hidden) antecedents added up; sometimes this was not speculation, but something approaching straight-forward description, plain enough for all but the patient to see.

Since I have urged that mentalism does not belong in the center of a map of the concept of thinking, the reader might suppose that I would reserve that spot for an avowal orientation. After all, I have argued that behavioral dispositions are inherently subject to differing interpreta-

tions, and that avowals typically settle disputes. However, I do not believe that an avowal orientation can provide the heart (if there must be a heart) either. I do not believe that a person's preparedness to avow "I think x" can be essential to an analysis of "he thinks x." My reasons for this resemble Wittgenstein's "private language argument" discussed in the earlier section on objections to mentalism. Recall that Wittgenstein argued that in order for the use of an expression ("think") to be taught there must be adequate criterion for determining that the expression applies, and for deciding whether the child has mastered its use. This line of reasoning provided an attack on mentalism. By the same token, it cannot be *essential* that a child be prepared to avow "I think x" before we can correctly make the attribution "you think x," for then we could *never* be correct in the cases we were using as our basis of instruction.

I conclude, then, that if there is a center to the map of the concept of thinking, it is charted by the behavioral orientation; the mentalist and avowal orientations may encompass rich territories, but they are more peripheral, and we do well to first familiarize ourselves with what is more central, the behavioral dispositions which unpack our language of attributing thought.

I would hope that what I have said might be found relevant to what psychologists call "attributional" theory. Having endorsed a behavioral orientation to the meaning of "thinks" I would want to underscore what I remarked parenthetically early in this paper, that what I have called a behavioral orientation is by no means to be confused with what psychologists call "behaviorism." In the first place, I would regard myself as a "soft nosed metaphysical behaviorist": I would want to include (but not stress) events like silent speech, which happen to be (contingently) private and yet could, in principle, be replaced by public events, among the "behavior" which I regard as relevant to the concept of thinking. In the second place, in much of the psychological literature on attribution theory and on behaviorism there is a commitment to a *causal* perspective which I regard as opaque and fundamentally muddled. I have explained my sympathies for an existentialist viewpoint elsewhere (Russell, 1978b, 1979). I do believe that human behavior is explainable and predictable; I do not believe that this entails anything at all about causality, nor do I believe that what is profound in attribution theory need have any association with the causal language with which it has become ill-wed. I have hinted in this essay that the ordinary language notion of thinking is rooted in phenomena accessible to the social scientist, and is a useful and valuable tool to employ in the work of explaining human actions. I should add my opinion that a fuller understanding of the concept of thinking would require a much larger map which pointed out interdependencies between this concept and several others in the same family, such as "action," "intention," "desire," and "choose." When such a map is drawn I believe it will show that the concept of causality never had any coherent place on it, but this claim takes me well

beyond the scope of the present paper.

References

- Austin, J. L. *How to do things with words*. Cambridge, Massachusetts: Harvard University Press, 1962.
- Russell, J. M. Psychotherapy and quasi-performative speech. *Behaviorism*, 1973, 1, 77-86.
- Russell, J. M. Saying, feeling, and self-deception. *Behaviorism*, 1978, 6, 27-43. (a)
- Russell, J. M. Sartre, therapy, and expanding the concept of responsibility. *The American Journal of Psychoanalysis*, 1978, Vol. XXXVIII, no. 3, 259-269. (b)
- Russell, J. M. Sartre's theory of sexuality. *Journal of Humanistic Psychology*, 1979, 19, 35-45.
- Ryle, G. *The concept of mind*. New York: Barnes and Noble, 1949.
- Sesonske, A. Saying, being and freedom of speech. *Philosophy and Rhetoric*, 1968, 1, 25-37.
- Wittgenstein, L. *The blue and the brown books*. London, England: Basil Blackwell, 1958. (a)
- Wittgenstein, L. *The philosophical investigations*. New York: The MacMillan Company, 1958. (b)