

ABSTRACT  
**IMAGES OF IDENTITY**  
***In Search of Modes of Presentation***

Ruth Garrett Millikan  
University of Connecticut

There are many alternative ways that a mind or brain might represent that two of its representations were of the same object or property, the "Strawson" model, the "duplicates" model, the "synchrony" model, the "Christmas lights" model, the "anaphor" model, and so forth. I first discuss what would *constitute* that a mind or brain was using one of these systems of identity marking rather than another. I then discuss devastating effects that adopting the Strawson model has on the notion that there are such things as modes of presentation in thought. Next I argue that Evans' idea that there are "dynamic Fregean thoughts" has exactly the same implications. I argue further that all of the other models of thought discussed earlier are in fact isomorphic to the Strawson model. A search for the source of these difficulties reveals the classical notion of modes of presentation as resting on two assumptions, neither of which I recommend. It depends on denying that the way the mind reacts to or understands the thoughts or ideas that it harbors has any bearing on their intentional contents. And it depends on an internalist view of thought content, in particular, on denying that the natural informational content carried or potentially carried by a thought has any bearing on its intentional content.

**IMAGES OF IDENTITY**  
***In Search of Modes of Presentation***  
RUTH GARRETT MILLIKAN

1. Introduction: mental sameness markers

In *Subject and Predicate in Logic and Grammar*, Strawson offers "a picture or model" of what happens when a man learns that two things formerly thought to be separate are in fact one and the same. "We are to picture a [knowledge] map, as it were" on which all those individuals the man knows of are represented by dots, and the predicates the man knows to apply to each are written in lines emanating from these dots or, if the predicate is relational, lines joining two dots.

Now when [he man] receives what is for him new information...he incorporates [this] by ...making an alteration on his knowledge map [for example,] he draws a further line between two dots. But when it is an identity-statement containing two names from which he receives new information, he adds no further lines. He has at least enough lines already; at least enough lines and certainly one too many dots. So what he does is to eliminate one dot of the two, at the same time transferring to the remaining one...all those lines and names which attach to the

eliminated dot. (Strawson 1974, pp. 54-55.)

On Strawson's picture, the identity of a particular is represented in the mind by the identity of another particular. So long as you haven't made any mistakes, everything you know about your mother is attached to the same *particular* mental representation of your mother, the same token. Your understanding that all these facts are facts about the same woman consists in the representations of the logical predicates of each of these facts being attached to numerically the same "dot" in your mind or brain. Call this the "Strawson model" of how identity or sameness is thought.

A more familiar model pictures thoughts each as a separate sentence token in a mental language. On this model the identity of a particular is represented by the identity of a mental word type rather than the identity of a token or particular. What Strawson would model using a single dot and two lines, a language of thought model renders as two different sentence tokens containing a word type in common, say, <Tom is married> and <Tom is harried>. Generalizing this to any system in which sameness is represented by duplication of form, we can speak of the "duplicates model" of how identity is thought.

Another model taken from language is the "equals sign" model of how identity is thought. Here a second marker riding piggyback on the duplicates marker joins *examples* of two different types with a mental equals sign. The effect of this "identity belief" is that all tokens of both exemplified types are treated as representing the same.

An absorbing contemporary discussion among cognitive neurologists concerns the "binding problem". Neurological evidence indicates that various kinds of sensory information arriving from the same object, such as information about form, color, and direction of motion, are not processed in the same area of the brain but filtered through "widely disseminated feature-detecting neurons located even in different areas or cerebral hemispheres" (Engel 1993). How then is it represented that these various features belong to the same object, and not to entirely different objects merely co-present in the perceptual field? One hypothesis is that synchronous spiking in neural firing patterns on a millisecond time scale indicates which sets of neurons are responding to the same object. Roughly, cells that fire together purport to talk about the same object; identity is represented by synchrony. If identity might be represented

---

1. I will move back and forth between idioms appropriate to traditional thinking about minds, and idioms more appropriate to thinking about brains, on the assumption that the structural forms we will be comparing are abstract enough to justify this. Theories of thought inevitably proceed on the assumption that there are abstract analogies between how thoughts work and how more mundane things work or might work. Think, for example, of Plato's *Theaetetus* with its mind that talks to itself, its wax imprints and its birds, or to the classical tradition that ideas are "like" their causes or that ideas are "associated" in the mind, and so forth. My talk about "the Strawson model", "the Christmas lights model", "the synchrony model", and so forth, should be understood in the same spirit.

this way in perception, why not also in thought? Call this the "synchrony model" of how identity is thought.

Connectionist explorations suggest as a crude model that units representing the same object might be strongly connected so that they tend to be activated together like Christmas tree lights on the same string. Then a certain causal connection would represent identity. Call this the "Christmas lights model" of how identity is thought.

Anaphoric pronouns, which occur in all natural languages, suggest a model according to which each representation of the same object bears some kind of pointing relation to prior representations of that same object. Call this the "anaphor model" of how identity is thought.

And so forth.

Now it is strongly emphasized in the Fregean tradition that representing the same thing twice, representing it once and then again, must be carefully distinguished from representing it *as being* the same thing again. If someone represents Mark Twain to herself and then represents Samuel Clemens, she represents the same thing twice, but it does not follow that she represents or understands *that these are* the same. Reflecting on the above models, we see that we should be equally careful to distinguish between representing the same thing twice *in the same way* and representing it *as being* the same thing again. If someone represents Mark Twain to herself in a certain way and then represents Mark Twain again in exactly the same way—if she duplicates a representation of Mark Twain—it does not follow that she represents or understands that the two referents are the same. To assume this would beg the question how the mind or brain represents identity. It would assume that *duplication* is used by the mind/brain as a sameness marker prior to any evidence that this is the case. There is no reason to suppose in advance that whenever sames represent sames, *sameness* must represent *sameness*.

Suppose, for example, that you observe the same individual apple in exactly the same context from exactly the same angle under exactly the same lighting conditions on two different occasions, and make exactly the same perceptual and cognitive response to it each time. *Merely as such*, this fact does not *constitute* that you recognize the apple as being the same apple again. On the other hand, there are relations other than duplication among percepts that mark object identity across time straightaway for the human perceiver, namely, the right *continuities* in perceived place and time. Given the right continuities, one's perception may be of an object as being the same one over a period of perceptual tracking despite its apparently changing in every one of its observed nonrelational properties. This sort of tracking of an object, say, with the eyes, does not involve *repeating* some particular way of recognizing the object over and over, but is more like following its path.

What I have been saying about mental markers for identity of individuals also applies, of course, to markers for other kinds of sameness. Just as representing the same individual twice in the same way is not representing it *as* the same individual, representing the same property twice in the same way is not representing it *as* the same property—not unless duplication happens to be what the system uses as its sameness marker for properties. We cannot assume without evidence, for example,

that whenever the same color, shape, or distance are represented in perception the same way twice, once here, say, and once there, one *ipso facto* recognizes these properties as *being* the same. We cannot assume without argument that duplication of any response, whether physical, perceptual or cognitive, is a sameness marker. A system of thought might also use different sameness markers for different kinds of identities. As Strawson described his own model, the sameness markers for predicates were not what I have called "Strawson markers", but rather duplicates markers.

In this paper, first I address the question what would *constitute* that a mind or a brain was using one method of marking sameness rather than another. Then, using the Strawson model as my example, I discuss certain difficulties that emerge when, beginning with models less familiar than those taken from language, we try to reconstruct the Fregean view that intentional attitudes are characterized in part by particular "modes of presentation" through which the referents of their constituent terms are grasped, and which modes can be differentiated (individuated) by rational reflection. A search for the source of these difficulties reveals the classical notion of modes of presentation as resting on two assumptions, neither of which I recommend. It depends on denying that the way the mind reacts to or understands the thoughts or ideas that it harbors has any bearing on their intentional contents. And it depends on an internalist view of thought content, in particular, on denying that the natural informational content carried or potentially carried by a thought has any bearing on its intentional content.

## 2. Locating the sameness markers in thought

Suppose that the cognitive neurologist—or God—looks down into the mind/brain with an eye to deciphering which of its various states or events are the ones representing identities. How is the neurologist or God to tell, given a mind in motion, *how* it is thinking identities?

First, we might ask, on what evidence do neurologists *in fact* suppose that synchrony may be the brain's marker of identity for perceived objects? The evidence they give is that synchrony is in fact found (in monkeys and cats) among cells responding to the various properties of numerically the same visually presented object. At least within the more accessible visually involved layers of the brain, information about one and only one individual object feeds into one synchrony, information about others into other synchronies.

Generalizing the neurologist's method, evidence for some feature being the sameness marker used by a system might be that information derived from the same thing in the environment systematically showed up marked by this marker. Thus, evidence that Strawson's model was right would be that all and only structures bearing information derived from numerically the same environmental source showed up attached to numerically the same something-or-other in the brain or mind; evidence that the duplicates model was the right model would be that all and only structures bearing information derived from numerically the same source showed up attached to structures alike by some specifiable principle of likeness, and so forth.

There is an obvious problem with this method, nor has it escaped notice by the neurologists. Synchrony among neuronal firings caused by the same object may be only a byproduct of the brain's perceptual activities. That these neurons fire

synchronously may have no connection with any cognitive work done by the brain. That a bit of natural information about sameness of source resides in the brain does not prove that the brain *uses* or *understands* this information, any more than the presence of natural information in the sky carried by black clouds proves that the sky thinks it will rain. Compare the hypothesis that certain neurons in the visual system are "feature detectors". The circuitry that produces firing of such neurons may seem to be intricately specialized to support this function, but the final proof must demonstrate that the firings are *used as* feature detectors, that is, that the information collected by them actually guides the organism to take account of the features apparently "detected". Similarly for whatever is found in the brain or mind that appears to be a sameness marker. What the neurologists would like to show is that synchrony is not just a natural indicator of sameness, but is effective in guiding thought and action to take account of the indicated sameness. It appears then that we must start further back. We must ask what is involved in *using* a marker as a sameness marker. What does a mind have to do in order to manifest understanding of its own sameness markers?

Begin by asking why it should matter to any organism whether or not various pieces of information that it has acquired are about the same object or about different objects. Why does an organism need to have sameness markers in perception or thought? Suppose that I perceive that  $\alpha$  is orange and that  $\beta$  is round and that  $\gamma$  smells sweet and that  $\delta$  is fist-sized and that  $\epsilon$  is within reach. Why does it matter whether  $\alpha=\beta$ , or whether  $\delta=\epsilon$ , and so forth. Because if  $\alpha=\beta=\gamma=\delta=\epsilon$ , but only then, probably this is a reachable orange, hence can provide me nourishment. Only by using these various bits of information *together* can this understanding be reached, but these bits can be used together legitimately only if they all carry information about the same.

Moving from the level of perception to the level of thought, suppose that I believe that A is smaller than B and that C is smaller than D. Only if I also understand that  $B=C$ , can I make an inference: A is smaller than D. Suppose that I believe that Cicero is bald but that Tully is not. Only if I also understand that  $Cicero=Tully$  can I discover that I am involved in a contradiction. In this sort of manner, every mediate inference, every recognition of a contradiction, and nearly everything learned from perception depends upon recognition of the identity of various items as one and the same thing having multiple properties or, if a property, multiple subjects. Consider how perception and thought are joined to guide action. Suppose that I wish to congratulate  $\alpha$  on his engagement and that I see that  $\beta$  is in the lounge talking to  $\gamma$ . This seeing will be of no use to me unless I grasp whether  $\alpha=\beta$  or  $\alpha=\gamma$ . Consider learning. Suppose baby has noticed that A scolded her when she cried but that B, C and D kindly picked her up. Whether (rightly or wrongly) she learns anything from this will depend on which if any of these four she takes to be the same person again. And so forth.

The lesson is simple but very important. When one piece of information is joined with another to yield more information, or to yield action, or to yield inductive learning, generalization, or practical learning, one or more "middle terms" are always involved. Some middle terms are predicative (*A is smaller than B* and *B is smaller than C...*) and some are propositional rather than denotative (if P then Q, and P, therefore Q) but at least one repeated element is invariably involved. It is for the sake of such movements in thought that recognition of sameness in thought content is required. I will refer to all

such movements as "mediate inferences": "practical mediate inferences", "theoretical mediate inferences, "perceptual mediate inferences", and so forth.

For mediate inference to occur, there must be recognition or positing of sameness in referential content of a middle term occurring in each of two premises. I propose, indeed, that we put this more strongly:

*For a thinker to recognize sameness of reference of terms JUST IS for the thinker to be disposed to pair these as a middle term in making mediate inferences.*

That will do for a first pass over the phenomenon of recognizing sameness.

A second pass must take account of the fact that where valid mediate inferences are made, this must result from some kind of indication from the premises themselves where sameness of reference is occurring. It must result from a system of sameness marking in thought, perhaps using Strawson-style markers, and/or duplicates markers, and/or Christmas light markers, and so forth.

*What makes these markers sameness markers is that the cognitive system uses these markers to control the mediate inferences it makes.*

Derivatively, then, the mere occurrence of an appropriate sameness marker bridging two thoughts or ideas can count as an "understanding" that the marked ideas are ideas of the same. It is, as it were, a "first act understanding of sameness", where a "second act understanding" involves the actual process of mediate inference as guided by these markers.

A third pass over the question what it is for sameness of reference to be recognized in thought should take error into account. Under unfavorable conditions, even simple perceptual identification tasks can be mismanaged. For example, there is a way of crossing your fingers so that the identity markings that bridge between tactile and visual percepts become mixed. The finger one sees being touched does not seem to be the same as the finger one feels being touched. *Conceptual* responses to the data of sense are far more tenuously correlated with affairs in the world than are perceptual responses. Failure to mark sameness correctly in thought is quite common. We often fail to recognize a thing, or we confuse two things together, say, mistaking Jim for Bill or failing to distinguish between mass and weight. Consider, then, a mediate inference that is made over two premises containing information in fact derived from different sources. The premises do not reflect information concerning the same thing, and as a result, let us suppose, the conclusion arrived at is false. Should such an erroneous move count as a mistake in inference? Or should it count merely a mistake in data collection and labeling? The difficulty is that which internal moves should count as valid inferences would seem to depend on how sameness of origin is marked during data collection. But how sameness of origin should be marked during data collection depends on what sameness markers the inferencing systems will recognize. There will be nothing wrong, for example, with representing two different objects with exactly the same mental form so long as mental form—broadly, duplication—is not the identity marker. Does it follow that which structures really are the sameness markers is well defined only for a system that never makes mistakes?

If we are naturalists, we can refer to evolutionary design on this sort of question (Millikan 1984, 1993b). There will be ways that our perceptual-cognitive systems

worked when they operated such as to be selected for, that is, a way or ways that they were "designed" to mark and to recognize sameness. In general principle, what these ways are should be no harder for us to distinguish than, say, how the human eye is designed to work, despite the existence of many eyes that function poorly.

### 3. Naive Strawson-model modes of presentation

Suppose that our minds/brains used Strawson markers for marking identity. Keeping clearly in mind that our project at this stage is neither Strawson exegesis nor Frege exegesis, let us ask what, on this model, would correspond most closely to the Frege-inspired notion that the same object can be thought of by a thinker under various different "modes of presentation". Gareth Evans tells us that different modes of presentation are, just, different ways of thinking of an object (e.g., Evans 1982, section 1.4). Suppose that we take this statement completely naively. On a Strawson model it appears that, so long as we always recognized it when we were receiving information about an object we already knew something about, each of us would end up having only *one* way of thinking about each object. No matter what attributes the Strawson-style cognitive system thinks of an object as having, as long as it does not fail in the task of reidentifying it always thinks of the object the same way, with the same dot. Two modes of presentation of the same might occur, for example, as the system collected information about a person seen in the distance prior to recognizing them, or about a person being discussed by gossipers before finding out whom they were talking about. But this sort of situation is usually temporary, either because the person seen or discussed is soon identified, or if they are not, because information collected about an unknown person is easily forgotten. For example, we do not usually retain memories of people we pass on the street if we don't recognize them. On this model, it would usually be so that all your beliefs concerning the same object were beliefs entertained under precisely the same mode of presentation.

On this naive reading of "modes of presentation", moreover, no two people could think of an object under the same mode of presentation. To do so they would have to have numerically the same dot in their heads! On a Strawson model, there is no kind of similarity between two minds, either in internal features or in external relations, that would constitute their thinking of the same "in the same way. There might be relevant similarities between the ways you and I think of a thing, for example, we might have exactly the same beliefs about a thing, associate with it all the same identifying descriptions and so forth, but on this interpretation this would not bring us any closer to thinking of it under the same mode of presentation.

Interpreted this way, "modes of presentation" obviously would bear scant resemblance to Fregean senses, the very first job of which was to correspond to shared meanings of words and sentences in public languages. For example, Frege supposed that the very same senses are grasped first by the speaker and then the hearer when communication is effected through language. Further, on this model the different identifying descriptions that you attach to the dot representing a given man are not different ways of thinking of him, but merely various things you know about him, some of which *might* sometime come in handy in helping to reidentify him as the source of some incoming information—no more. Correspondingly, the difference between

various kinds of referring expressions—descriptions vs proper names vs indexicals—would not parallel a difference between various kinds of thoughts. On this model there would be, for example, no indexical thoughts or ideas, although there would, of course, be times when the thinker used perceptual tracking abilities to collect bits of incoming information together next to the same dot in his head.

And, of course, sentences expressing nontrivial identities could not be analyzed Frege's way given this model. Accordingly, Strawson's description of the semantics of identity sentences in (Strawson 1974) did differ radically from Frege's. The public meaning of the identity sentence is neither given by its truth conditions nor does it correspond to a particular sharable thought. It concerns what the sentence conventionally *does* to hearers' heads. What it does is not to impart information but to change the mental vocabulary, altering the mental representational system. As such, its function is different, in one important sense, for every hearer. Both the affected dots and, barring weird coincidences, the information in the structures attached to these dots, will be different for each hearer.

Perhaps most important of all, on this model, should the thinker make a mistake in identification the result will be the creation of an equivocal mode of presentation, one that has two objects at once. Nor will the one who "grasps" this equivocal mode of presentation have access to this fact. Suppose that you are confused about the identity of Xavier, having mixed him up with John, so that whenever you meet either you store the information gathered next to the very same dot. Which man does this dot represent? If systematic misidentifications occurred, or if misidentifications were frequent and random, it seems that a dot's reference might scatter over a multitude of objects, hence, reasonably be considered quite empty, or that a dot might undergo massive yet invisible shifts in reference.

A corollary would be that negative identity sentences would have no determinate meaning for a person. For example, on this model you have no separate ideas <Cicero> and <Tully>, nor even <the man called 'Cicero'> and <the man called 'Tully'>. Your way of thinking of the referent of each of these four public terms merges them irretrievably together. Suppose then that a historian now informs you that there has been, in fact, an unaccountable confusion among philosophers and that Cicero was *not* in fact Tully. How are you to understand this negative identity claim? What you've got in your head is one dot, attached to which is a variety of (presumed) information, including the information <...is called "Cicero"> and <...is called "Tully">. But how will you divide the rest of the information into two piles? This could only be accomplished through a major job of reconstruction, as you tried to remember or to guess how you had acquired each separate bit of information, hence from which of these men it was most likely to have originated.

It is hard to imagine anything further from Frege's intention than these various results. What has gone wrong? I will soon tease apart several strands that are woven together to produce the peculiarities of this "naive" Strawson-inspired image of "modes of presentation", and I will try to articulate the underlying principles that divide it from Frege's own vision. But first, there is another interpretation possible of what modes of presentation might be for a mind that used Strawson markers.



#### 4. Strawson-style modes of presentation as ways of recognizing

In our "naive" image above, Strawson's dots are taken to be modes of presentation because they are "ways of thinking of things", a phrase most easily interpreted in this context to mean, kinds of mental representations of things. In interpreting modes of presentation this way, we departed from an important strand within, anyway, the contemporary *neo*-Fregean tradition. Gareth Evans equates the way one is thinking about an object with the way in which the object is identified (Evans 1982 p. 82, McDowell's formulation for Evans). Similarly,<sup>2</sup> Dummett takes Fregean sense to be a method or procedure for determining a *Bedeutung*, paradigmatically, for determining the presence of the *Bedeutung* (e.g., Dummett 1973, pp. 95ff). Evans and Dummett agree, for example, that grasp of a particular way of recognizing a referent when encountered in perception corresponds to a mode of presentation of the referent. Now the Strawson image of sameness marking seems to pry apart the way one thinks of a thing from the various ways one knows to recognize it. Perhaps, then, if we identify modes of presentation with the latter instead of the former we will find them to be more as Frege intended.

Suppose then we take modes of presentation, on the Strawson model, to be not ways of *thinking* about a thing but ways of *identifying* it, in particular, ways that a thinker knows to recognize incoming information, arriving via perception or inference, as being about a thing. On this reading, it seems that a person might well grasp not just one but many modes of presentation for a given object. Also, different people might well grasp the same mode of presentation, for they might be able to recognize the object in the same way. Moreover, perhaps understanding an expression that refers to an object is, just, grasping a particular way to identify the object, so that the meanings of referring expressions will correspond to modes of presentation (though the contrary would not always be true.) Certainly there is some plausibility to this for the case of identifying descriptions: one can sometimes recognize that something is the object in question by first knowing that it fits a certain description. How this might be so with public-language demonstratives is harder and with proper names (since Kripke's *Naming and Necessity*) perhaps hardest of all, but certainly there are philosophers who have supposed something like this to be true. The results thus look better at first than on the "naive" interpretation.

But trouble is not far away. On the Strawson model, the terms in the various *beliefs* that a person has will not be characterized by determinate modes of presentation. Characteristically, each dot will be coordinated with multiple ways of identifying, multiple ways that the thinker would be able to recognize incoming information about that referent. But the various pieces of information attached to a given dot are not associated with any one of these ways more than another. For although each bit of information may have found its way to the dot by just one path of recognition, the Strawson system keeps no record of which information entered by which path. Besides, on this model modes of presentation are not supposed to be just

---

2. Similarly enough that is. Evans is at pains to distinguish his views from Dummett's here, but not in ways that affect what is at issue for us.

ways a thing has historically been recognized by the thinker, but ways s/he knows to recognize it. But if the terms of a thinker's beliefs are not characterized by determinate single modes of presentation but by many modes at once, and if these various modes should be such that in fact they do not all determine the same object, then, as before, it seems that the thinker's thought might be in all innocence equivocal.

The interaction of Strawson's image of sameness marking with Frege's vision of modes of presentation thus seems to yield strikingly unFregean results. Soon I will argue that these results are not merely an artifact of the Strawson model; they follow given any model of sameness marking. Strawson's way of marking identity merely spotlights features easily overlooked when employing other models of sameness marking. But first, I want to point to a way in which these apparently radical results are not far removed from at least one contemporary neoFregean vision. Evans' view that there are "dynamic Fregean thoughts", if pushed to its limit, yields exactly the same results.

#### 5. *Evans' "dynamic Fregean thoughts"*

Evans (Evans 1981, 1982 p. 194ff) proposed that when you are tracking an object perceptually, say, keeping it in view as it moves and you move, if you continue to believe over this period of time that the perceived object has a certain property, this should not be considered to be a sequence of similar beliefs that you have, but a single belief that persists over time. You continue to think of the object under the same mode of presentation, as long, that is, as you have indeed kept track of it. Evans calls this sort of thought a "dynamic Fregean thought", and he says that in such cases the relevant "*way of thinking of an object*" is a "*way of keeping track of an object*" (p. 196). Now if you do not merely persist in the same belief about the object over the tracking period, but continue to collect new information about it from perception, noting, say, its way of moving, what it looks like from the back, what it sounds like, how large it is and so forth, presumably this will not change the fact that you continue to think of it under the same mode of presentation, as long as you don't lose track. Note the isomorphism with the Strawsonian analysis of the same tracking event: you continue to keep many old predicates attached to the dot while you funnel in various new bits of information to attach to the very same dot.

And what if you should unknowingly lose track of the object? You thought it was one little minnow—you named him "Primus"—that nibbled first your toe then your ankle, but there were actually two. In that case, Evans claims, "we have not a case of misidentification but a case where the subject has no thought at all" (Evans 1982, p.176). For in the absence of "an ability to keep track" of the object, "it is not possible for a subject to have a thought about an object in this kind of situation at all" (p. 195).

But an ability is not, in general, something one either has or has not. Most abilities come in degrees. One of my surest abilities is my ability to walk, but there are still times when I trip. Hence there seems another response possible for Evans. You do have an ability to keep track of things like minnows, only this time you tripped. Hence this particular dynamic mode of presentation of yours is indeed part of a thought, but the thought happens to be equivocal. It hovers between the two minnows, presenting both as if one. True, Evans is wedded to "Russell's principle"—"that a

subject cannot make a judgment about something unless he knows which object his judgment is about"— and he interprets this to mean that the subject "has a capacity to distinguish the object of his thought from all other things" (Evans 1982, p. 89). But Evans gives no argument that I can tell for the soundness of this principle used this particular way.

I am happy to agree that if a mode of presentation were sufficiently equivocal, not just mixing little minnow Primus with Secundus, but also rolling in, say, Sextus and Octavius, and, indeed, also a large random sample of other minnows in the school, it would be odd to consider it as determining a thought of any minnow at all. It should probably be considered "a case where the subject has no thought [anyway, of minnows as individuals] at all". Some ability to track them individually seems necessary to having thoughts of individual minnows. This is parallel to the result I earlier suggested we would get on the naive Strawson-inspired model if enough mistakes in identifying were made: the resulting thoughts would be empty. But I have never heard an *argument* anywhere that *no equivocation at all* is ever possible in thought.

Whatever one decides about that, however, surely the case of error-infected naive Strawson-inspired modes of presentation and the case of error-infected Evans-inspired dynamic modes of presentation must be decided in the same way, for the parallel is exact. The parallel can be shown, indeed, to be a structural identity. Consider the dynamic mode of presentation involved as you perceptually track a person, Kate, to whom you have just been introduced at a party. For a brief moment—not much longer, suppose, than a saccade—you divert your eyes to the face of a friend, but immediately pick up Kate's face again. Then a large fat man, excusing himself, passes between you and Kate, but again you immediately pick up the track. Looking at Kate and hearing her voice, you perceive these as having the same source, as locating the same person. Now Kate passes for a moment into another room, but you continue to hear her voice—though of course there are spaces between the words—and she soon emerges again. By now she is beginning both to look and to sound quite familiar, so that after stepping outside for a moment, you immediately find her again. The time interval was longer this time than between her words, but short enough for her voice still to be "in your ears". Compare this, for example, to the way a bloodhound tracks a person by smell, at moments losing but then picking the scent up again. One should not think of the bloodhound as merely *repeating* a particular way of recognizing the person over and over as the scent is lost and regained. Now suppose that Kate looks and sounds familiar also an hour later and then a day later when you meet her again, first in the lobby, then on the street. Probably you would not have recognized her, however, had you met her in Singapore—in some radically disjoint context. Similarly, Evans tells us in his chapter on recognition that even though by using one's recognitional ability alone one might not be able to tell a certain sheep one is thinking of from every other sheep in the world, still, because one can keep track of the neighborhood in which the sheep is likely to be and one can also keep track of where oneself is, one can maintain an ability to reidentify the sheep (Evans 1982 8.3). Further now, suppose that Kate's name has become familiar, and as more time goes by you often pick up information about her from friends. Again, you usually know which "Kate"

they are talking about from the context.<sup>3</sup>

When did you stop tracking Kate? When did you stop following her spoor, the trail she left of ambient energy structures bombarding your sensory surfaces? When did the original mode in which she was presented to you come to an end?

A dynamic mode of presentation that never came to an end would be, functionally, exactly the same as a "naive" Strawson-style mode of presentation. Each of the peculiar, distinctly unFregean traits that I have described for the latter modes would characterize the former as well. Whether Evans' dynamic modes really differ from Strawsonian modes in function depends, then, on whether a clear principle of individuation—of sameness and difference—could be drawn for abilities to track. When did you leave off one "ability to track" and start using another "ability to track", or some different kind of ability to "know which object you are thinking about", as you collected information over time about Kate?

#### 6. Does it actually matter how sameness is marked?

What I would like to understand, is what the basic principles are that distinguish the vision of thought generated from the Strawson image of sameness marking from Frege's original vision of thoughts as exemplifying modes of presentation. The first conclusion I will reach is that, surprisingly, the way the Strawson markers mark identity plays no role in determining this difference. Strawson markers merely highlight a general feature present in all other models as well.

Consider, for example, duplicates markers. How will a system consisting, say, of mental sentences and that uses only duplicates markers come to realize that Cicero is Tully? It must put all the Cicero and Tully information into sentences using the same name, either <Cicero> or <Tully>, choose which. Just as one of the dots has to go on the Strawson model, one of the shapes has to go on the duplicates model. So if it should turn out that Cicero is not in fact Tully, whichever mental name got chosen will be equivocal, nor will the news that Cicero is not in fact Tully represent, for the system, any definite instructions for separating the information again into two piles. Duplicates markers do not differ in any way from Strawson markers in function.

Frege might be interpreted as having supposed that the mind uses, in part, a duplicates system of sameness marking. For although senses were not supposed to be psychological entities, *graspings* of them surely are dated, psychological occurrences, and Frege seems to have held that it is awarenesses of duplicate graspings-of-sense that keep us from contradiction and govern the performance of rational mediate inferences. If so, we might note, this constitutes a substantial psychological claim: a perverse deity might have made our minds otherwise. Imagine, for example, the same sense coming into mental view twice simultaneously as subject term of contradictory judgments, but the demon has determined that only synchronously vibrating viewings of the same sense will move the mind to recognize sameness of reference. One thing that is clear, however, is that Frege did not view the result of an identity judgment to be the elimination of one of the two kinds of thoughts

---

3. Support for the line taken in this paragraph is in (Millikan forthcoming).

involved, nor did he suppose any thoughts were equivocal.

Now Frege also took there to be two kinds of identity judgments, the "informative" ones such as "Cicero is Tully", and those such as "Tully is Tully" which are not. We might suspect, then, that it is the introduction of this second way of marking identity, used for identities not known *a priori* to hold, that allows the Fregean thinker to identify referents without merging his thoughts of them. This second identity marker, we suppose, functions like a mental equals sign. It marks two thoughts as being thoughts of the same, not by merging or destroying either, but simply by flagging them for use together in mediate inference.

Yet if we think carefully about the effects of an equals marker on the system that understands it, it becomes puzzling how they would differ from the results of a Strawson marker. After all, exactly *which* similarities determine two word tokens to be examples of the same word type clearly depends on the practices of language users—here, the practices of the cognitive system using the tokens. In the case of a public language, these practices can be pretty disorderly, the rules can be quite "disjunctive", and can include many exceptions. For example, in English, contrasting pronunciations of "schedule" (s-k-e-dule vs. sh-e-dule) count as tokens of the same word type while exactly the same contrast between the pronunciations of "skeet" and "sheet" or "skin" and "shin" produces different word types. And of course these practices may change over time. Surely the same should be true for mental word types. Whatever the individual mind/brain *treats* as the same mental word again IS the same word again.

Nor is there any reason why mental typing should not *evolve* in an individual mind or brain. In this particular respect, the situation is not like that of a public language, where there are typing conventions laid down in the public domain, prior to a particular person's use. For mental language, nothing corresponds to these conventions. But what effect are we to imagine mental <Cicero = Tully> to have if not, precisely, that it changes the mind's dispositions to mental typing? Henceforth, mental <Cicero> and mental <Tully> will behave as representational equals. They become the same mental word. But if this is so, the mental equals marker behaves exactly like a Strawson marker. It merges two thought types into one, threatening equivocation in thought, and doling out to each thinker just one mode of presentation per object.

We must conclude, I think, that the peculiar effect of the Strawson markers was on us, on our understanding, not on the operation of the cognitive systems modeled. Systems that use Strawson markers understand identities by *explicitly* changing their mental vocabularies, replacing two representations with one. Systems that use equals markers do exactly the same thing but implicitly, changing merely the typing rules for their mental vocabularies, that is, merely the *functions* of the symbol "shapes"<sup>4</sup>.

---

4. It is possible to imagine a system in which a log was kept of the various changes made in the representational system as identity judgments were made, and a log kept of the inferences that had turned on these supposed identities. Compare the way modern word processing programs keep track of the last hundred or so commands carried out. Then if a mistake were discovered, the "undo" button could be pressed until the system was returned to the point of the original false identity judgment. Different

### 7. The first Fregean assumption

How then does the Fregean avoid the Strawson image with its threat of equivocation in thought and its frugal offer of just one mode of presentation per object? By assuming that how a thought *functions* has no effect on its content. By assuming that how the mind *understands* its thoughts is irrelevant to their significance. On the contrary, I have implicitly assumed throughout that use does affect representational value. I assumed that what marks referential sameness in thought is whatever the cognitive systems *read* as marking referential sameness, or what they are designed to read as marking sameness. I assumed that if thought tokens are marked to function as representing the same, this will affect their representational value. In particular, if this marking conflicts with other factors relevant to representational value, say, with the informational content of the tokens so marked or other ways their associated referents may be determined, then there will be equivocation in content. Sameness is represented yet different things are represented.

The Fregean view assumes, on the contrary, that insertion of a sameness marker, hence change in the use of the marked terms, has no bearing on content. Placing a mental equals sign between <Cicero> and <Tully> has no effect on the representational value of either, even if Cicero is not in fact Tully. Similarly, if duplicated thoughts are in fact thoughts of the same, each token of <Cicero> referring again always to the same rather than acting, say, like the English word "he", this depends in no way on the fact that duplication is what is *read* by the mind as marking identity. Thought typing is determined independently of thought use.

This Fregean assumption is equivalent, I believe, to the assumption that thoughts are not mental representations. For we cannot suppose that a representation could be a *mental* representation, a representation *for mind*, yet that its representational value was independent of its effect upon mind. And of course Frege did not hold that thoughts are mental representations. Senses are abstract entities that bear their contents quite independently of how a mind "grasps" them. The conclusion that

---

pieces of information that had become attached to the same Strawson dot at different times would then have different statuses in case of emergency. Indeed, might one say that they represented predicates attached to the same subject but under different modes of presentation? How many of the various purposes of Frege's modes of presentation could differences of this sort serve? I have not explored these questions because I think such a model is completely lacking in psychological plausibility. Imagine keeping such a log on all the times you have ever reidentified or made inferences about your husband or mother! It is true that were I seriously to suppose, say, that Mark Twain was not Samuel Clemens, I might indeed have some idea how to guess which of my beliefs about him should be attached to which name. Certain facts would cohere with Twain's role as an author, others perhaps with his role as public speaker or builder of the Twain house in Hartford. But this untangling would certainly not be done on the basis of a memory of when and in what order I had discovered or inferred what about Twain.

classical Fregean modes of presentation are not compatible with a representational theory of mind is not, as such, a criticism of Frege. But it may serve as a warning to some would-be neoFregeans.

. The second Fregean assumption

According to Frege there are informative and also uninformative identity claims. Uninformative identities are so called because they do not inform us of anything not already immediately known *a priori*. Presumably they also cannot be false. Frege is not supposing that there might be false identities that we cannot help but affirm. The Fregean senses that uninformative identities equate function psychologically as would thoughts marked with duplicate markers. But it is not for this reason that duplicate graspings of Fregean senses always have the same content. Not that such a reason is impossible on other views. On a representationalist view, sameness markings do force both marked thoughts to refer to the same thing, hence if the markings are wrong, forces both to refer equivocally. On a Fregean view, however, the referents of duplicate thoughts are entirely separately determined. That the thinker identifies the referents as one and the same is in no way responsible for them being the same. What is the guarantee, then, that the referents of duplicate thoughts actually ARE the same? Or if what you mean by "duplicate thoughts" *includes* that they have the same referent, what is to guarantee that the mind that grasps two thoughts can *tell* whether these thoughts are indeed duplicates? How can there be uninformative identities that are at the same time certain really to be identities and not merely false appearances of identity?

This line of questioning highlights the internalist assumption built into the Fregean position.<sup>5</sup> What is duplicated when "the very same thought" is repeated has to be something which is at the same time (1) compelled always to bring with it the same referent and (2) capable of being unmistakably known by the mind (while the mind is doubly entertaining it) as being the very same thought. That, I take it, is one role of a Fregean sense: it always determines the same referent regardless of the context, the grasper (understander), or the use, and its identity is transparent to mind.<sup>6</sup> More generally, that which completely determines the referent must be exactly the selfsame as that which, when duplicated, constitutes an understanding of sameness. Otherwise the understanding might not be veridical. It follows that whatever determines the reference must be entirely internal to mind. Reference cannot be affected at all by, say, the external causes of thoughts, or their natural informational content, for there can be no certain *a priori* mark to prove that the external causes or the informational contents

---

5 Compare, for instance, (Millikan 1993 chapter 14: "White Queen Psychology").

6. If it were transparent to mind, then that one can not think a contradiction about a thing while thinking of it under the same mode of presentation could not be criterial of sameness of sense, nor could it be assumed that uninformative identities were never false identities. (If it is not Frege's position that identity of sense is transparent to mind, it is certainly what many have thought was his position. The purpose here is not, of course, Frege exegesis but clarification of where certain compatibilities of position lie.)

of two thoughts are really the same.

We can make the situation clearer by redirecting our attention to the role of modes of presentation in the Fregean tradition. Letting the question rest whether ways of recognizing a thing correspond to ways of thinking of it, let us look again at "ways of identifying" or "recognizing" as candidates for modes of presentation. If we follow Evans in taking there to be "dynamic Fregean thoughts", there is a major problem (section V above) concerning criteria of identity for "abilities to keep track" of an object. But even if we don't follow Evans, there is a similar and more general problem about individuating "ways of recognizing". There is, that is, if we also adopt an externalist account of "ways of recognizing" a thing, taking these to be, not ways of holding an object up before the mind, but ways of knowing when one is receiving information about an object, either through sensory input or inference.

Modes of presentation are supposed to be individuated in such a way that the rational subject can always tell one mode from another. Also, each mode of presentation presents always the same object. That is why the rational subject need not fall into contradiction nor make invalid mediate inferences. Suppose then that modes of presentation are ways of recognizing things as just described. Then there must be methods of counting out at these ways of recognizing such that (1) each cannot fail to net always the same object and (2) this fact is guaranteed *a priori*. But of course there are no such guarantees. There is no such thing as an ability to reidentify that isn't fallible. That is the famous Achilles heel of verificationism.

Perceptual evidence never guarantees its sources. Perceptual tracking is of course fallible. There is never an *a priori* guarantee that one has kept track. The same is true of recognitional abilities. You may know, for example, literally hundreds of ways to identify each member of your immediate family, some of which ways—a long look full into your spouse's face in full daylight, for example—may (barring removal of your brain to a vat) actually be infallible. But if that is so, it is because the world, not anything in your mind, is constructed so as to make it so. It is because there is not in fact any other person in the world who looks just like that in the face (and no one *actually* able, and desirous of, putting your brain in a vat)—a convenient fact but not one guaranteed by *a priori* reflection. Recognition using identifying descriptions is, intrinsically, even more fallible. First, that the description is unique is always a contingent fact. There might always be, within limits of discernability, two tallest or two oldest, for example, so that neither is really tallest or really oldest. And one can always make a mistake about which one IS tallest or oldest because one perceives wrongly, or because one infers wrongly, or because one is informed wrongly by others. True, it has seemed to many that an identifying definite description is the surest sort of tool one could use to make fixed what one was thinking about. But the job that must be done by a method of recognition for incoming information is not to fix a thing before the mind. It is to effect *actual* reidentifications, to direct *actual* incoming bits of information about the same to a focus so that they will interact with one another in inference. Used for this purpose, most definite descriptions are of severely limited value.

But if no ways of recognizing are sure always to capture the same, then there are no modes of presentation that are sure to be univocal, hence no modes under which it might not be *correct*, leave aside whether it would be "rational", both to affirm



and to deny. The externalist now has a choice. Either no two pieces of information about an object are ever believed under the same mode of presentation, in which case no valid mediate inferences can ever be made. Or else all pieces of information about an object are grasped under the same, possibly equivocal, mode of presentation, so long, that is, as they are identified as being about the same.

I conclude that the notion of modes of presentation is not a useful notion for any theorist who has befriended mental representations, nor for any who has befriended externalism about thought content. The reason is that there is no principled way to individuate modes of presentation on these views, or<sup>7</sup> to achieve any of the various effects for the sake of which Frege introduced them.

RUTH GARRETT MILLIKAN

*Philosophy Department  
University of Connecticut  
Storrs CT 06269-2054 USA  
e-mail Millikan@UConnVM.UConn.edu*

---

7. I am grateful to Michael Martin for catching an error of substance and both to him and to William Lycan for requesting footnote #4.

## BIBLIOGRAPHY

- Engel, A.K. 1993: "Temporal coding in the visual system; implications for the formation of representational states in neuronal networks". Presented at the conference on "Naturalistic Approaches to Representation and Meaning", Zentrum für interdisziplinäre Forshung, Universität Bielefeld, December 1-3.
- Evans, G. 1981: "Understanding Demonstratives", in Herman Parrett and Jacques Bouveresse, *Meaning and Understanding*. Berlin and New York: De Gruyter.
- 1982: *The Varieties of Reference*. Oxford: Clarendon Press.
- 1985: *Collected papers*. Oxford: Clarendon Press.
- Millikan, R.G. 1993a: "On Mental Orthography", in Bo Dahlbom, ed., *Dennett and His Critics*, Oxford: Blackwell, pp. 97-123.
- 1993b: *White Queen Psychology*. Cambridge: MIT Press.
- 1994: "On Unclear and Indistinct Ideas", in J. Tomberlin, ed., *Philosophical Perspectives* vol 8, Atascadero. CA: Ridgeview Publishing, pp. 75-100.
- (forthcoming): "A Common Structure for Concepts of Individuals, Stuffs and Basic Kinds: More Mama, More Milk and More Mouse". *Behavioral and Brain Sciences*.
- Strawson, P.F. 1974: *Subject and Predicate in Logic and Grammar*. London: Methuen & Co Ltd.