Loosing the Word-Concept Tie

Ruth Garrett Millikan
University of Connecticut
ruth.millikan@uconn.edu

Abstract

Sainsbury and Tye propose that, in the case of names and other simple extensional terms, we should substitute for Frege's second level of *content* — for his senses — a second level of meaning vehicle — words in the language of thought. I agree. They also offer a theory of atomic concept reference — their 'originalist' theory — which implies that people knowing the same word have the 'same concept.' This I reject, arguing for a symmetrical rather than an originalist theory of concept reference, claiming that individual concept are posessed only by individual people. Concepts are classified rather than identified across different people.

1. Introduction

Sainsbury and Tye intend their theory of concept identity to account for the main data that led Frege to his theory of sense while avoiding well known problems with that theory, first about the senses of indexicals and demonstratives, second, the problem raised by Mates's examples seeming to imply that no two words can have the same sense. As I understand it, S&T's solution has two important parts. First is the proposal that, in the case of names and other simple extensional terms, we should substitute for Frege's second level of *content* — for his senses — a second level of meaning vehicle — words in the language of thought. If Peggy thinks that groundhogs are larger than woodchucks, the content of her thought is indeed contradictory, but this contradiction does not show up in the vehicles of her thought any more than it shows up on paper, so she cannot be accused of being irrational. Similarly should she happen to think that Cicero is Caesar rather than Tully. When she learns that Hesperus is Phosphorus, rather than learning that two senses capture the same referent, she merely learns to attach all the same mental predicates to her mental 'Hesperus' that she attaches to her mental 'Phosphorus.' Further, although S&T don't themselves make this move, perhaps to believe the assertion that Homer did or did not exist is merely to proceed to engage or to disengage one's mental term 'Homer' from interacting with one's ordinary thoughts and actions in the usual way. And Mates's problem appears to be solved as well. Pablo can wonder whether everyone who believes that the Greeks are the Greeks believes that the Greeks are the Hellenes because these beliefs may contain different words in these people's mental vocabularies. S&T take mental words to be atomic concepts. The thesis implies, then, that different atomic concepts can sometimes have the same referent or extension.

I like this sort of view of 'the Frege data.' Indeed, as spelled out only this far, it accords with my own position as outlined in (Millikan 1984 Chs. 9, 12[1]; 1993 Ch. 14; 2000 Ch. 6; 2004 Ch.7; 2005 Ch. 3). However it requires to be supported — as S&T proceed to do — with a theory of atomic concept reference that explains how different atomic concepts, separate mental Millian names for objects, properties and so forth, can sometimes have identical contents. S&T's title, 'An Originalist Theory of Concepts,' refers to their proposal on this second matter, and it is this originalist proposal that I will discuss. I think that it leads them to a problematic position on, as they call it, 'the individuation of atomic concepts' and on the relation of mental words to public language words. I will explain the problems as I see them and then briefly propose a non-originalist position on atomic concept reference, on the individuation of atomic concepts, and on the relation between mental words and public words that I believe gives better support to their (and my own) proposed analysis of 'the Frege data.'

2. Originalism

The originalist theory begins by adopting some version or other[2] of Kripke's 'picture' (from Naming and Necessity) of how the reference of a proper name is determined, then generalizes this to

[1] In Millikan 1984, 'intensions' served as a (very bad) term for what, in later work (e.g., Millikan 2000) I called 'conceptions,' roughly, ways of recognizing manifestations of a thing. The 'Frege cases' were dealt with using the notion of 'language bound intensions,' that is, abilities to recognize ***linguistic*** manifestations of a thing.

[2] Major portions of S&T's paper are presented 'for illustrative purposes' only, which makes it hard to pin things down. In some cases it may be that my exposition of their position and also my criticisms should be read as 'for illustrative purposes' as well.

include all public terms expressing atomic concepts. Someone has used a term to refer to something; others then copy the term, deferring their reference to whatever the first user intended, and so forth in a chain leading up to current uses of the term. Call these 'originalist Millian terms.' Current people may use an originalist Millian term without any more understanding of what it is a term for than bare syntax and context supply. Originalist Millian terms are individuated by history, not physical form or user understanding. Terms sounded or spelled the same way but that do not historically converge on the same original meaning event are different originalist terms. S&T emphasize that different originalist Millian terms can have the same meaning without current people knowing it. It seems to follow as well that same-sounding originalist Millian terms might have different meanings without people knowing it, a theme to which I will return.

Next, originalism apparently applies the above strategy also to mental terms or concepts introduced by a thinker for his own use. S&T say, for example, that various occasions on which a speaker thinks 'that' to himself may or may not 'involve distinct specific demonstrative concepts.' There are 'originating specific that concepts' and 'non-originating specific that concepts,' the difference being 'inferred from the speaker's intentions and reactions'. For example, when the thinker introduces an originating specific that concept 'he has no inclination to bring forward the information' that was associated with any earlier specific that concept use, whereas 'non-originating cases... are marked by deference to previous uses.' Presumably various different 'originating specific that concepts' when remembered and deferred to later on, will have acquired the status of separate mental Millian names, just as homonyms in a public language, if understood as such, will normally correspond to separate mental terms. Different 'specific that concepts' will turn into different mental words unless one of them 'defers' to a previous one and the thinker is 'inclined' to 'bring forward the information' associated with the previous one to apply to it.

This seems to imply a general sort of model for Millian concept introduction without language. One observes something, gives it a new mental name, and when one later (correctly or incorrectly) takes something else one is observing to be the same thing again, one deferentially applies the same mental word again, or borrows it, also carrying over previously acquired information about this thing to the new occasion. Thus mental Millian terms, concepts, that have been introduced without language will also be individuated by history, their meanings deferring to their original introductions. S&T explicitly demure from speculation on how children acquire concepts prior to language, yet they seem to have supplied us with a rough sort of model for this anyway. And just as in the case of publicly originalist Millian terms, it also seems clear how, due to failure to reidentify, a person might harbor separate privately originalist Millian mental terms for the same thing without knowing it. But assuming that the very same mental term is never accidentally introduced twice, originalism seems to imply that there could not be a privately originalist mental Millian term with a double meaning — a privately derived mental homonym caused by misidentification— because *only the very first use has a determining effect on meaning*. That is what makes the theory originalist. I will return to this also.

Next, S&T's originalism describes the relation of public Millian terms to mental Millian terms. Following Putnam, Burge and others, when a concept is acquired by hearing or reading a new public language term, a new mental word is introduced whose meaning the thinker borrows from that of the public term. Thus the origin of the public language term directly supplies the meaning of the new mental word. Further, if you and I derive mental words in this way from the same public term, then we will have 'the same concept,' for 'the key positive thesis of originalism' is that 'concepts are individuated by their originating use.' The view is thus consonant with the (ubiquitously held) supposition that we can talk about '*THE* concept dog' or '*THE* concept water,' that we can do 'conceptual analysis' by trying to figure out what a certain public-language word means, and so forth. Many people have the same concepts, these concepts corresponding to shared words. I think we should be clear, however, that this view could not imply that 'having the same concept' involves having in mind the same mental word orthographically speaking, as it were. However the neural orthography of thought is supposed to go, we have no reason to suppose that my mental word corresponding to the English word 'dog' looks or feels like yours. either to us or to from neurologist's point of view. Similarly, your mental terms corresponding to 'bolt' of cloth and to 'bolt' the door are not the same.

Finally, what does the originalist think happens when a person has a privately acquired Millian name for something first and only later encounters a public language term for it? A first possibility would seem to be that he doesn't recognize that he already has a mental word for this something and adds a redundant mental word to his mental vocabulary. He then thinks there are two things when there really is

only one. But what if he does recognize the identity between mental term meaning and public language term meaning?

To remain stoutly originalist, it seems to me that S&T should maintain that he would use the public term, both when understanding and when speaking, by deferring to the meaning of his earlier mental term. For surely he will have an 'inclination to bring forward the information' that was associated with earlier uses of the privately acquired mental term to bear on uses connected with the public term. Indeed, recalling that we should not suppose that the orthography of the public term must determine the orthography of the mental term with which it coordinates, the natural supposition is that he will employ the privately acquired mental term as his correlate for the public one. Why confuse himself by suddenly changing his mental vocabulary? But then, since originalism holds that 'concepts are individuated by their originating use,' it would follow that there were many names and other words in English for which your corresponding concepts and mine were not the same, for they will have had different origins, having originated for one or the other or both of us as privately acquired mental words.

According to S&T, however. '[i]n our view, the concepts infants form on their own are typically supplanted by public concepts when they become full members of their surrounding linguistic community.' Presumably they would say the same about the concepts you or I form on our own — 'that man,' 'that building' — before we find out what the community calls these things. Apparently these privately originated concepts just disappear. Pure originalism suddenly gives way before lingua-centrism: one's language of thought, in so far as it coordinates with one's public language, is nothing whatever but internalized public language. And we are back with the comfortable assumption that we can talk unambiguously about 'the concept dog,' and that 'conceptual analysis' is trying to find out what a word in the public language means.

3. Worries about Originalism

I am going to worry about three aspects of S&T's originalism. First, I will question the asymmetry of originalism. I see no reason to suppose that when I make (what I take to be) a later sighting (as it were) of something I had also sighted earlier, the thought that I retain of that something is any more of the first sighted thing than of the second. A parallel symmetry applies to the meanings of public words as they are used over time. Crucial to understanding both cases is the possibility of misidentifiction, a topic not addressed by A&T.

Second, I will question S&T's lingua-centrism. I see no reason to suppose that the public language simply displaces prior concepts. A result of these first two questions will be to question that speakers using the same word generally back it up with 'the same concept.' Millian concepts, I will suggest, are indeed pretty well modeled as mental words, but mental words are things that we do not share. We each have our own. On the other hand, we can of course classify mental words, concepts. Classification by content, by reference, is often the most useful when referring to other people's concepts, but since a person can harbor referentially redundant concepts, it is not invariably best. Other kinds of classification can be used too, including classification according to public words through which the concept's referent is sometimes identified by the thinker. But uses of 'same concept' to refer across thinkers are always only classificatory, not individuative.

Third, S&T discuss what makes one concept the same as another and they discuss at some length whether or not you have to 'know the content' of a concept you have. But they don't tell us how a concept is originally related to its content and just what it would be for one to 'know' that content. Rather, they explicitly defer to future progress in philosophy of language and mind to solve these problems. I will attempt to fill these gaps, though for Millian empirical concepts only. The result will uphold S&T's claim that the 'Frege data' are explained by turning to the level of mental words rather than to a second level of content, but I will oppose their originalism.

4. Symmetry

First then, symmetry. Suppose that I encounter a new person whose name I do not learn. I reidentify them after having turned my back, then after having returned from the next room, then on the front porch, then again on the street next day and so forth, until they have become a familiar figure. Suppose that the moment I first met them has completely disappeared from my mind. In order to talk to you about this person I will have to offer a description, or I may have to point and say 'that person,' but for myself, presumably, they are defined not by any particular description or that thought. They have, for me, a certain Millian mental name. But who that Millian name refers to would not seem to be determined by any one encounter I have had with this person more than another. Every use I make of my mental name for them 'defers' to all of the others. Even supposing that I do remember the original meeting, why would my

mental name be somehow partial to it? Surely, for example, I will bring forward information gathered on every past meeting to any new meeting quite impartially. Similarly for a mental name I originate for a species of wild flower, or for a tune, or for a certain font, or a symbol, or a style of architecture, or a color that sticks in my mind, and that I reidentify on various occasions but for which I lack a public-language term

But now the possibility of misidentification enters (not discussed by S&T). Suppose there have been occasions on which I mistook someone else for this person, and I bring information from those misleading occasions along to current encounters also? Will my mental name now refer to the misidentified person or persons and to the original person equally?

Yes, a bit, but probably not at all equally. It would surely be rare for very many misidentifications to occur, the marginal fuzz, if any, left on the content of the mental name being quite insignificant. For example, suppose that fuzz happened to contain my very first encounter when I first coined my mental name but after that I consistently identified and reidentified a different person under that mental name. The referent of my mental name would then seem to be the second person, information, if any, retained from my very first meeting having turned into misinformation! (Would originalism really be a more plausible call in this case?) On the other hand, suppose that I hopelessly confuse a pair of twin brothers, thinking there is only one, not two. I form just one mental name that brings forward information collected about both brothers equally whenever I use the name. Then, yes, my concept is hopelessly equivocal, being of one brother mixed equally with the other. A theory that allows a thinker to have two mental names for one entity should allow one name for two entities as well. It should recognize the possibility of one mental word standing in, confusedly, for two things for the simple reason that occasionally this actually happens. A long meaning-rationalist tradition, spun out over many generations of philosophers, rested on the assumption that only public terms can be equivocal, any apparent confusions in concepts, in mere thought, being detectable a priori and analyzable, ultimately, into mistaken judgments. But I know of no argument that was ever given for this assumption, hence of no reason to continue the tradition.

5. Lingua-centrism

Suppose I have a well-used mental name for a person but I don't know their public name. I then discover that their public name is 'Suzy.' Knowing this, I have now acquired a way of recognizing incoming information about Suzy that is clothed in a new medium. Before I could recognize infusion of Suzy information into the ambient energies impinging on me only when it came in a form enabling me to see her, perhaps to recognize her voice, to see or hear things that were reliable natural signs of her presence, and so forth. Now, given common sense, caution and sensitivity to context, I can recognize, with reasonable accuracy, various bits of information about her that reach me through language. (In Millikan 2004 Chs. 3-6, I explain how language, when it carries intentional information, normally carries natural information that coincides.) Indeed, through this medium I might chance to pick up more information about Suzy in a dozen minutes than I had accumulated before in a dozen years. But there seems no reason to suppose that I will change my mental name for her. Indeed, it may happen that I retain all the new information but soon forget her public name. Or although I am thinking of her and want to tell you what I am thinking, at the moment I can't recall her name. Had I learned her name was Suzy and been told various things about her first, picking up more information directly from her presence here and there later on, there still seems no reason to suppose an asymmetry in determining the reference of my mental name for her between the referent of 'Suzy' and ways I have recognized her other times. Surely we would need an argument for asymmetry here, and we would want to know how this asymmetry shows up in thought or behavior.

Recalling again that mental names can't just be mental copies of public names, should I later learn that Suzy is also called 'Mrs. Tomkins,' I will freely bring forth, combine, all the information I have on 'Suzy' with what I hear about 'Mrs Tomkins' quite symmetrically, continuing to harbor only one mental term for Suzy, only one concept of her. Similarly, my Hesperus concept can surely defer to my Phosphorus concept and vice versa, thus merging into one and the same concept, being thought with the same mental name. I have a daughter with at least a half dozen nicknames, some used by some people some by others. Surely I don't have a half dozen different concepts of her corresponding to these! All of these considerations would seem to apply as well in the case of Millian names for flower species, for tunes, for fonts, for symbols, for a style of architecture, for a color, and so forth, and also for my concept of, my mental name for, woodchucks/groundhogs. (Otherwise, notice, the bilingual must be supposed to have double concepts of almost everything, two complete mental vocabularies.)

6. Correcting Kripke's 'picture' of the reference of a name

These reflections may carry us back to wondering how the reference of the public name 'Suzy' was itself determined. Gareth Evans (1973) alerted us long ago that something was amiss with Kripke's 'picture' of how the reference of a proper name is determined, since proper names — his example was 'Madagascar' — sometimes change their referents. They can also be equivocal, as with 'Homer', who was probably a compilation of many Greek poets over many generations, or 'St Patrick,' who has been said to be two Irish bishops that got mixed together (O'Rahilly 1942). Other Millian terms[3] have changed their meanings too; 'jade' is one example (*The Encyclopedia Britannica* 1967), 'meat' (as S&T themselves point out) is another. Clearly, something is needed here more or other than bare originalism.

I have described what I take that other to be at some length in my (2010). The idea, with some reservations, is that if a term has survived, continued to be reproduced, because, in context, it has carried information about X sufficiently often, then its referent is X. But, of course, the referent of a term can sometimes be vague or equivocal, even though we individuate terms, as S&T do, by lineage rather than physical form. For our purpose here let me offer just a 'picture' (in the spirit of Kripke) of public-meaning stasis versus public-meaning change.

A Millian term M is introduced and correctly understood by original hearers as a name for m, these hearers proceeding to use M with that understanding. That is, they start using the name M along with other ways they already had or that they acquire for identifying incoming information about m and they use M for disseminating information about m. Later hearers sometimes first pick up information about m merely by recognizing repeats of the term M, sometimes later recognizing information about m in other ways too and sometimes not. There may also come along hearers who misunderstand M, identifying its referent with things other than m. If enough of them identify the referent of M with, say, n, the term M may become equivocal between referring to m and referring to n, and if, eventually, nearly everyone identifies the referent of M with n, then M will have come to refer to n. In-between things can happen too, leaving the term M equivocal or somewhat equivocal or very vague but, of course, its meaning may instead remain sharp and static. There could also be points at which the term M branches into several lineages, several distinct senses, that begin to proliferate independently ((Millikan 1984 Ch. 4; 2005 Chs. 3, 10).

7. What is an empirical concept?

Now to fill in the gaps. What is an empirical concept, how is it connected to its content, and what is it for the thinker to know its content?

The simplest animals govern their behaviors entirely by reflexive or tropistic reactions to proximal stimulations of the organism. If behaviors are to be governed also by reference to distal environmental contingencies, the cognitive engineering problems to be solved suddenly increase dramatically. The difficulty is that the same distal object or property may have any of numberless different proximal impacts on the organism's sensory surfaces, depending on distance, direction, partial occlusion, mediating circumstances and energies, disturbances or interferences of kinds that are innumerable even in principle. Consider, for example, how many different kinds of stimulations impinging on your senses may result in your recognizing incoming information about a member of your family — as seen from the front, from the back, from different distances at different angles, sitting, standing, walking in the distance, in daylight, at dusk, by lamplight, in spotted sunlight, by their voice, by their voice over the phone, by their voice across a crowded noisy room, by their clothes, by any of thousands of descriptions, by their handwriting, by what they are doing, and so forth. The same difficulty is involved in recognizing any empirically known property that an object may have. Psychologists have been hard at work for many years trying to figure out how the perceptual systems manage to achieve color constancy or shape constancy or depth constancy or, given that phonemes are not sounds but rather vocal tract gestures, how phoneme constancy is achieved, and so forth. Apart from being directly perceived, incoming information about objects, properties, and so forth, is often identified by hearing names, or descriptions, or with the help of inference. Indeed, everything you know about a thing may, under some circumstances, help you in identifying it, or help, as importantly, to prevent you from misidentifying it. No, that can't really be gold, because the price tag is too low. Yes, that must have been James; he's the one who always wears a beret. The stuff's gone green; there must be copper in it (Quine). Concepts caught up in scientific theories are no different in this way from ordinary empirical concepts. The more ways discovered to identify the presence of the same object or to measure the same property by observation and/or

[3] For a defense of the claim that many, indeed probably most, simple empirical terms are Millian or lacking in intension, see my 2010.

inference, the more secure is the confidence in the real existence of that object or property.

The ability to recognize the same distal object, property, relation etc. as being the same through a wide variety of manifestation to one's senses is needed in order to collect over time both factual and practical/behavioral knowledge concerning it and to know when to apply that knowledge. For an animal that makes mediate inferences, theoretical or practical, it is needed to supply nonequivocal 'middle terms,' (in a generalized sense) for inference. (For example, the 'middle term' for modes ponens is the antecedent of the conditional.) All of these uses require that one collect to a focus information that concerns the same thing but that has been scattered from its original source in many directions through diverse and often bumpy media, bringing it to bear as a unit on action and thought. It involves, using S&T's words, that one 'bring forward the information' about a thing that has been gathered through many diverse media to bear jointly when needed. Having such an ability with regard to an object, property, relation or real kind, etc., and knowing ways to use it is, I suggest, having an empirical concept of that thing. Having such an ability can also be modeled, for many purposes, as having a single mental word which one uses to represent that thing in the context of numerous different intentional attitudes that concern it.

On this view of empirical concepts, a concept is an individual thing— an individual ability that comes into existence at a time or over a time and goes out of existence at or over another time and that belongs to an individual person or animal. I have concepts and you have completely other concepts, though many of them may be concepts of the same things. That our concepts are of the same things does not mean that we use the same methods of identification for these things. I may recognize Suzy by her looks and voice while you can only recognize her by her name. Consider Helen Keller's concept of red or her concepts of thunder and lightning. Also, it is easy to see how a person can have two concepts of the same without realizing it, how it might be news to someone that Hesperus was Phosphorus or that groundhogs are woodchucks or that the familiar man on the commuter train is Representative Hic Slick. It is equally easy to see how a person might have one concept of two people, having mixed them together in her mind. And it is easy to see how a person might seem to herself to have a concept that was not in fact of anything, hence that was not, strictly speaking, a concept at all. This could happen because she was attempting to identify through an empty word, handed down to her along with fancy descriptions, say, about someone coming down the chimney on Christmas eve, wearing red and white, chuckling 'Ho ho ho' and so forth. And it can also happen in other ways, to be discussed below.

8. What determines the content of an empirical concept ?

Causal and informational theories of empirical concept content face two related problems. The first is misidentification (Fodor's 'disjunction problem'). If the content of a concept is taken to be what regularly causes it's tokening, or what it's tokening carries natural information about, what are we to say about misidentifications — cases in which mental HORSE tokens are caused by seeing cows on dark nights? The second concerns the specificity of the relevant cause or information that is supposed to determine content. Every tokening of a mental term has numerous causes and carries many layers of natural information. For example, typically a mental representation 'skips over (or 'sees through') the intermediate links in the causal chain in order to represent ... its more distant causal antecedents' (Dretske 1981, p. 158). How are the relevant cause and the relevant level of distality determined? These issues are entwined, so my discussion of how my suggested theory (which is neither a causal nor, strictly, an informational theory) avoids them is one.

My suggestion is that having an empirical concept involves the ability (and the ability to use the fruits of this ability) to reiterate a term in one's mental vocabulary when natural information about the same thing arrives again at one's sensory surfaces, arriving, typically, in any of a variety of different forms, perhaps including the impact on one's senses of certain public language utterances or inscriptions. We can dramatize the idea this way. When I acquire a new belief containing the mental word **cow** this is usually the result of having encountered new natural information about a cow or cows, or having made an inference from prior beliefs about cows. Also, and crucially, this pattern results not accidentally but from the successful employment of specialized cognitive structures, hence of abilities. Abilities (very roughly - please see Millikan 2000, Ch. 4) are dispositions that have been designed by evolution, or by any of various kinds of learning, for certain definite purposes, purposes which they achieve if the mechanisms supporting them are not damaged and -- much more problematic -- if they are operating under Normal conditions and with Normal input. Normal input and Normal conditions are the kind that allowed these mechanisms to perform their tasks in past cases that lead to their being selected or retained.

Having thus imported a distinction between Normal and abNormal cognitive functioning it may

seem that a solution to the error or 'disjunction' problem should be at hand. Suppose then that we add the idea, very close to explicit in S&T's paper, that one of the things the cognitive systems are designed to do is to reiterate a mental term when they receive information that concerns the very same thing again and to carry information about that thing forward accordingly. A mistake is made when a mental term is tokened in response to information about something other than has produced its tokening before. — But in this way we run headlong into the specificity problem! Which of the many things the first token helped to carry natural information about are the cognitive systems supposed to be tracking when iterating the same term again? Suppose that they have already iterated mental **FIDO** over and over in response to natural signals that carried information about Fido. These signals will all also have carried information about dogs, about animals, about furry things, about four legged things, about warm things, about sentient things, and so forth.

What's needed here is to look in the direction of use, as well as origin. Empirical concepts, I have said, are needed for collecting together, over time, both factual and behavioral knowledge about their referents, to recognize when that knowledge is applicable again, to supply nonequivocal 'middle terms,' for inference. What various kinds of things the cognitive systems were designed, by evolution and learning, to track information about and to represent consistently depends on the specific uses to which that information is, by said design, to be put. In the simplest and commonest cases, an animal's perceptual/cognitive systems are either designed to be able to reidentify certain definite affordances, or designed to become designed to reidentify affordances, that is, designed to learn to reidentify affordances. To identify an affordance is to recognize, for use, a certain objective kind of object, property, event, situation or whatever, to which a definite kind of response will, under Normal conditions, result in a definite and useful outcome, or a predictable outcome that sometimes has a use. Again, Normal conditions, with a capital N, are conditions that determined such outcomes in past cases that accounted for the selection or maintenance of the perceptual/cognitive mechanisms involved. (The job, anyway one job, presumably, of classical and operant conditioning is that of locating and learning to reidentify objective affordances.[4])

Concepts that are developed originally for forming ordinary descriptive beliefs — call them 'theoretical concepts' — are not designed to have direct connections with any specific behaviors. Descriptive beliefs help to control behaviors only indirectly, after adding conation and inference. To fill in a story about how a system might work whose job is to develop and use theoretical concepts requires a fairly deep excursion into ontology. It requires trying to understand how the structure of the world determines there to be objective identities of various categorial kinds that developing perceptual/cognitive systems can search for, and it requires exhibiting ways that success or progress towards success in learning to reidentify can be indicated to a learning organism. These are projects that I have undertaken in previous work, especially in (Millikan 1984 Ch. 14 ff; 2000 entire; 2004 Chs. 9-10). The position that I have defended includes the claim that theoretical concepts are developed for their roles in subject-predicate judgment sensitive to negation, this structure necessarily reflecting certain skeletal structures in the world, one import of this being that concepts of individuals and of real kinds can only be developed along with and in relation to concepts of empirical properties and relations falling within logical contrary spaces that pertain to them. For example, the core of the concept of any individual or real kind must include a rudimentary grasp, not of *answers*, but of the kinds of *questions* that can sensibly be asked about it — how old? how tall? where born? what color hair? speaks what language? Male or female? — rather than, say, what valence? how numerous? what kind of government? how high on the Richter scale? who wrote it? how much genetic variance? Where on the periodic table? and so forth. This kind of grasp — grasp of what I have called a 'substance template' for a (primary or secondary[5]) substance to be reidentified — is needed both in learning to track that individual or kind and in diagnosing mistakes in tracking. Very much more needs to be said here, however, very very much more, as is recognized and attended to in the references given above.

---

[4] To think of the 'discrimination and generalization' traditionally said to be involved in operant conditioning as discriminating among and generalizing over proximal stimulations of the afferent nerves is a serious mistake. Indeed, behaviorists stopped making this mistake relatively early, though that did leave rather a large gap in behaviorist theory.

[5] Very very roughly in the Aristotelian sense.

9. What is it to know the content of one's own concept?

In 1993 I contributed a paper to the Aristotelian Society-Mind meetings called 'Knowing What I'm Thinking of.' It concerned Russell's claim, and Gareth Evans's, that 'it is scarcely conceivable that we can make a judgment or entertain a supposition without knowing what it is we are judging or supposing about' (Russell 1912, p. 58). I argued for a new interpretation of what it would be to know what one was thinking of, from which it followed that this kind of knowledge is a matter of degree and can, in fact, sometimes fail. Let me begin by repeating a paragraph from that essay.

> What seems to be yearned for in the notion that I must know which object my thought is about is a sort of confrontation of thought, on the one side, with the object bare, on the other, taking place, per impossible, within thought itself. Indeed, Russell's view is that exactly this sort of confrontation is possible — the object bare is part of the thought. But, despite contemporary hyperbole that speaks of thoughts that require real objects in order to be thoughts as `Russellian thoughts,' no thought actually consists in part of its object — any more than a mother, though she has to have a child to be a mother, consists in part of her child. The closest thing to the yearned-for ideal that actually makes some sense, I suggest, is a confrontation of one thought of an object with another thought of that same object, taking place within thought itself, and effecting a recognition of the sameness of the object. Putting this picturesquely, if you imagine the various thoughts that you have about, say, Saul Kripke, as a sort of story that you tell yourself using various thought tokens of him (including perhaps perceptual ... thoughts of him), then knowing who you are thinking of in this story corresponds to your ability to make what Strawson called 'story-relative identifications' of the person in the story (1959, p. 18). There is no way that you can cut through the stories that you tell yourself about Saul Kripke in order to tack them inside your mind directly onto Saul, in order to know in any more direct way than that who you are thinking of. (Millikan 1993, p. 96)

That is, an act of correct reidentification, a re-tokening of the same mental term in response to incoming information about the same thing again, is an act of grasping what one is thinking of with that mental term, an act of gasping the content of one's concept. More generally, we can say that having an ability to reidentify that which the concept is of is knowing the content...a kind of know-how.

But abilities come in all degrees. One can know how to do a thing under many and diverse different conditions using any of a variety of methods, or only under a few conditions using one or a few methods. One can have an extremely trustworthy ability, or an ability that sometimes tends to fail. (Even my ability to walk occasionally fails me, when I trip and fall.)

Both Russell and Gareth Evans affirmed 'Russell's principle' that thinking of something involves knowing what you are thinking of. Russell also held that you can't suppose yourself to have a thought when you have none, whereas Evans held that you can suppose yourself to have a thought when you have none, exactly because you may not realize that you don't have the ability to reidentify. S&T deny Russell's principle, holding that you can have even a perfectly good concept without knowing its content. This explains, they say, how you can have an empty mental name without knowing it. If you don't have to know its content to have a concept, it makes sense that you also don't have to know if it has a content. What follows on these matters from the position I have outlined above?

Having a healthy empirical concept involves having (something we can usefully model as) a mental term that one knows how to iterate in response to incoming information about some particular thing, thus bringing to one focus information that has been gathered about that thing in a way that can enable certain practical or theoretical uses. In healthy cases, one pretty much always knows what one is thinking of, indeed, knowing what one is thinking of, having the ability to reidentify that thing is, in the first instance, what a concept is for — its raison d'etre. But concepts do not always succeed. We need to understand what different kinds of failures are possible. (How we label these kinds of failure is of less importance.)

A concept that fails pretty completely, failing to track anything objective at all, might, I suppose, be considered not to be a 'concept' at all. This would fit with Gareth Evans's idea that some seeming thoughts are not really thoughts at all. On the other hand, such concepts will have been designed by perceptual/cognitive systems whose job was to make healthy concepts, so just as something designed to open cans is a can opener even if it is too dull or too badly made actually to open cans, perhaps it is sensible to consider a concept that is badly made, hence fails, to still be a concept. (Indeed it is difficult to think what else one might naturally call it.) Empty names can give rise to completely failed concepts,

names that do not carry information about anything at all, or not about anything in the right ontological category to fit their attempted or proposed uses. Perhaps the concept of the ether, once supposed to be the medium of light waves, failed completely, and also the concepts of the four humors once supposed to determine a person's disposition and general health. (Recall, in connection with failed scientific concepts, that the various methods of reidentifying or of trying to reidentify associated with a concept may include or even consist entirely of inferential moves.)

Besides the possibility of complete failure, concepts can be equivocal, mixing two things together, or they may split several ways. Simple conceptual equivocations, such as confusing Jane with Jill, confusing jadeite with nephrite or confusing weight with mass, mix substances that fall under the same substance template or properties that fall in contrary spaces that apply to the same sorts of substances. Deeper conceptual errors may mix ontological categories, or suppose ontological categories that don't exit. Perhaps the concept of phlogiston was of this sort.

How then are we to answer the question whether one can have a concept without knowing its content, whether one can think of a thing without knowing what one is thinking of? As is quite often true in philosophy, it's probably better not to answer the question at all. We should rest content with understanding the nature and complexity of the real possibilities rather than arguing about how these possibilities should be labeled.

References

Evans, G. 1973: 'The Causal Theory of Names'. *Aristotelian Society Supplementary Volume* xlvii, pp. 187–208.

——1982: *The Varieties of Reference*. Oxford: Clarendon Press.

Millikan, R. G. 2010: 'On Knowing the Meaning; with a Coda on Swampman'. *Mind* 119, no. 473, pp. 43-81.

—— 1984: *Language, Thought, and Other Biological Categories*. Cambridge MA: The MIT Press.

—— *1993: White Queen Psychology and Other Essays For Alice*. Cambridge MA: The MIT Press 1993.

—— :1993. 'Knowing What I'm Thinking Of'. *Aristotelian Society, Supplementary volume* 67, pp. 109-124.

—— *2000: On Clear and Confused Ideas: An Essay about Substance Concepts*, Cambridge UK: Cambridge University Press.

—— *2004: Varieties of Meaning; the 2002 Jean Nicod Lectures.* Cambridge MA: The MIT Press.

—— *2005: Language: A Biological Model.* Oxford: Oxford University Press.

O'Rahilly, T. F. 1942: 'The Two St. Patricks'. Thomas Francis O'Rahilly papers, held at Queen's University, Belfast

Russell, B. 1912: *The Problems of Philosophy*. New York: H. Holt, Inc.

Strawson, P. F. 1959: *Individuals*. London: Methuen & Co.