

# 7 Cognitive Psychology and the Metaphysics of Meaning

Mark Johnston and Sarah-Jane Leslie

The philosopher's only resource is the analysis of concepts we already possess.

—Michael Dummett, *The Nature and Future of Philosophy* (2010)

In “Concepts, Analysis, Generics and the Canberra Plan” (Johnston and Leslie 2012), we distinguished the clusters of psychologically real heuristics that govern our use of terms—what cognitive psychologists often call “concepts”—from the philosophical notion of concepts as the meanings of terms, be they public terms, or mental terms in a supposed language of thought. Throughout what follows, meanings are understood as the semantic determinants of the extensions of the terms in question, and hence of the truth-conditions of the sentences that contain those terms.

Both the semantic and psychological notions of concepts are just fine, and not at all in competition, at least when understood as directed at different targets—the phi-concepts and the psi-concepts, as we propose to call them. But they are sometimes connected by an empirically discredited quasi-supernatural theory of our use of terms. The quasi-supernaturalist theory is that use is quite generally guided by *grasp of semantic meaning*, where this involves supposed occurrent causal influences on a speaker's psychological life by things neither physical nor mental, namely extension-determining “senses” or, in our terms, semantic meanings or phi-concepts.

In our view, there is no general explanatory factor that deserves the name “grasp of semantic meaning.” It is a misleading *façon de parler*. Aside from the cases where “grasp of semantic meaning” describes explicit knowledge of one or another of the comparatively few statements of any given language made true by meanings, knowledge to the effect that they are in fact so made true—cases such as “vixens are female foxes,” “a copse is a thicket of bushes,” and the like—the phrase “grasp of semantic meaning” is merely the description of an achievement, i.e. that of acquiring the ability to express one's thoughts by the use of the relevant terms in a given language,

and to understand thoughts so expressed. It is not a tenable account of some means by which the ability is realized.

Talk of semantic meaning has its proper home, not in a psychological theory of what guides use, but in the characterization of languages as abstract objects, where each such language is individuated by (a) a syntax and (b) a potentially infinite list of pairings of sentences of the language with their respective meanings. The potentially infinite lists can be characterized by exploiting a finite class of primitive subsentential expressions (as specified by the syntax), among them the names and simple predicates of the language, which are assigned their own meanings understood as encodings of their systematic contribution to the meanings of the sentences that contain them. If the language is potentially infinite, yet in principle learnable, then there will be a finite list of recursive rules for the combination of the meanings of subsentential and sentential expressions, rules mirroring the syntax of the language, which generate meanings for each of the potential infinity of sentences of the language. As a result of a specification of L's individuation conditions in terms of which meanings it pairs with which subsentential and sentential expressions, we can specify (1) truth-in-L understood as the statement of the truth conditions for the sentences of L and (2) analytic validity-in-L, which concerns just what argument patterns using terms of L are guaranteed to be truth-preserving solely in virtue of the meaning of those L-terms.

The question then becomes: what is it for a population to speak one of these abstractly individuated languages rather than another? In the jargon of the 1970s and 1980s: what is the actual language relation?<sup>1</sup> One standard account, which took various forms, was that a population speaks an abstractly specified language L if there are dispositions had by the speakers in the population to use the sentences of L in a variety of conventional speech acts such as assertions, directives, inquiries, etc., whose contents—*what* is asserted, commanded, inquired as to whether, etc.—are propositions with just the meanings paired with those sentences in the specification of the language L. There was an element of proper idealization in the determination of the relevant dispositions; they were to include ordinary speakers' dispositions to correct such usages as a result of deferring to distinguished speakers of the same language, e.g., in culture circles like ours, these include parents, teachers, and those scientifically in the know. So, to take Tyler Burge's familiar example, even someone who does not know that "arthritis" applies only to inflammation in the joints may speak a language in which this is nonetheless a condition on correct use of "arthritis," precisely because of that person's pattern of deference to distinguished speakers of the relevant part of the language, in this case doctors.

From the fact that a population, or group of speakers, achieves enough conventional coherence to speak a given language L, nothing very specific follows about just

<sup>1</sup> See David Lewis (1969, 1975), Stephen Schiffer (1972), Christopher Peacocke (1975), Martin Davies (1981), and Barry Taylor (1982).

how the dispositions to fall in with the conventionally required use of sentences are psychologically realized in the members of the population. After all, conventional conformity of use derives not just from the coincidence of the outputs of individuals' heuristics, but also from massive ongoing deference to the use of distinguished others around us; first our parents, then our teachers, then, in an age of science adoration, to those in the technical know.

In general, from the fact that two speakers at a given time speak the same language abstractly considered nothing very specific follows about what, as an empirical psychological matter, guides those individual speakers of the language in their use of terms in the language. Conformity in conventional use can cover a multitude of individual psychological differences among speakers in respect of the heuristics that generate that conformity of use. Likewise, one's heuristics can change over time, say as one learns more about a topic, without the meanings of one's terms for that topic changing.

In our view, the idea that what generally guides the speakers of L, the language they together speak, in their use of the expressions of L is a *grasp of the meanings* of such expressions, where meanings are then understood as the items that enter into the individuation of L, involves an odd mixing of explanatory levels, the psychological and the individuating. In familiar atheoretical parlance, to "grasp the meaning of a term" is just to know how to correctly use it in a range of not-too-far-<sup>Q</sup>ut actual and counterfactual circumstances. In its turn, that just involves having some effective heuristics or criteria for telling when one has, and when one has not, something to which the term applies. There is psychological reality behind talk of a speaker grasping such heuristics, but no general psychological reality behind talk of grasp of meaning. Again "grasping meaning" is the description of an achievement, not the description of some queer process by which this achievement is explained.

## 7.1 WHAT PSYCHOLOGY HAS TO OFFER

Thanks to cognitive psychology, we now know a great deal about just what our heuristics or criteria are like: they frequently take merely generic and prototypical forms; they often are guided by folk-scientific views about explanatory relations among the features of the thing at hand. The employment of such criteria frequently falls far short of even a tacit understanding of universal necessary and sufficient conditions for the application of the term. But that is the very thing which literally grasping the semantic meaning and being guided in one's use by such a grasp would provide. For in the semantic tradition, meaning is understood as determining sentential truth-conditions across all possible worlds, and in that vein the meaning of a subsentential expression—the concept expressed in the philosophical sense—is understood as determining the expression's extension and antiextension in every possible world.

Although the focus and the questions that drive psychological theorizing about concepts significantly deviate from the semantic tradition, this fact is partly obscured by an entrenched pattern of storytelling found in survey articles on the history of the psychology of concept use. Unfortunately, many summaries of the psychological literature on concepts begin with something called the "classical view," according to which subjects actually exploit represented necessary and sufficient conditions when applying most lexical concepts (i.e., a concept that is expressed by a single word). The standard illustration of the classical view is the concept *bachelor*, supposedly composed of the concepts *unmarried* and *male of marriageable age*, such that anything is a bachelor just in case it is an unmarried male of marriageable age. But, of course, not all concepts can be decomposable in this way; there must, on the classical view, be some or other stock of basic concepts, out of which all other concepts are ultimately composed. So "the classical view" encourages a picture of concept learning as combining such basic concepts to form complex ones. Bringing an item under a decomposable concept is then supposed to be a matter of checking whether the item satisfies the necessary and sufficient conditions specified by the decomposition. That is, the psychologist's target known as the "classical view" of the concept *bachelor* is not just the banal remark that it is true in virtue of meaning that someone is a bachelor if and only if he is an unmarried male of marriageable age. It is the specific empirical thesis that we actually use the concepts *unmarried* and *male of marriageable age* in deciding whether to count something as a bachelor. It is a thesis about the criteria we actually use, not, or not just, a thesis about the meanings of our terms.

Since the 1970s, the classical view, understood as a thesis about the criteria we actually use, has been quite roundly rejected. Much of the reason for its rejection has to do with the discovery by Eleanor Rosch and her colleagues of so-called typicality effects (e.g., Rosch 1973, 1978; Rosch and Mervis 1975). For many categories, some members of a category are perceived as being more typical examples of the category than others, and it turns out that how typical a category member is actually predicts a very wide range of experimental results. For example, people are quicker to categorize typical members, and are more confident and consistent in their categorization of typical members. When learning a novel concept, people learn to categorize the typical members first, and they learn the concept faster when presented with typical members in the learning phase. There are myriad other effects of typicality on language learning and use, on reasoning, and so on so forth. (For some very helpful reviews, see Laurence and Margolis 1999; Murphy 2002; Smith and Medin 1981.)

Hypothesizing that people are guided by, i.e. represent and exploit, necessary and sufficient conditions will not explain typicality effects. Knowing that something is a bachelor just in case it is unmarried and a male of marriageable age does not give any information about what makes for the typical (James Bond) versus the atypical bachelor (Pope Francis). The bachelors are all alike in respect of being unmarried males of marriageable age. Something else has to be posited to explain typicality

effects; but once this something else is recognized there may be no empirical reason to also posit the *actual exploitation* of known necessary and sufficient conditions. The classical view has fallen into disrepute because many investigators believe that precisely this has turned out to be true.

Consider, for example, an experiment conducted by Jerry Fodor and his collaborators, in which they asked whether one could find any differences in processing time that would indicate that one concept is composed in part by another. If, e.g., the concept *murder* is composed in part by the concept *kill* (as has been claimed), then it should take longer to process *murder* than *kill*, since processing the former involves processing the latter as a proper part. However, this prediction of the classical view is not borne out (Fodor et al. 1980).

Subsequently, Armstrong et al. (1983) found some typicality effects even for concepts such as *odd number*. This suggests that typicality effects are compatible with representing necessary and sufficient conditions. However, the crucial point is that in the case of concepts such as *odd number* we have independent reason for supposing that we represent necessary and sufficient conditions—namely that we (i.e., typical competent adult users of the term "odd number") can articulate what they are. There is no corresponding case to be made for the majority of concepts. Thus while the in-principle compatibility point concerning typicality effects and necessary and sufficient conditions is illustrated by Armstrong et al.'s findings, this in itself does not alter the fact that, for most concepts, we simply have no reason to suppose that ordinary, competent adults generally represent and exploit necessary and sufficient conditions suited to specify meanings. Indeed, like the Fodor result, the implication of Armstrong et al.'s findings is that even when speakers do know the semantic decomposition of a concept they are not exclusively using that decomposition as their heuristic for applying the related term.

Seeking to explain typicality effects, many psychologists were led to the prototype theory of the heuristics that make up our concepts. Prototypes can be understood as statistical functions over properties, which assign weights to features based on how likely a category member is to have that feature, or conversely, based on how likely something with that feature is to be a category member. There are a number of different proposals that fall under the heading of the prototype theory (see Murphy 2002, for an extensive review), but they all characteristically appeal to features that are in some way statistically related to category membership. For example, the prototype for *dog* might include features such as *barks*, *has four legs*, *has a tail*, *wears a collar*, and so on. These features are not candidates to figure in universal necessary and sufficient conditions since not all dogs have these features; an unfortunate creature can still be a dog even if it has three legs, no tail, no collar, and no bark. However, the basic idea behind prototype theory as illustrated by the concept *dog* is that if one is confronted with an animal and wishes to determine whether or not it is a dog, one will use this animal's features, or lack thereof, in a complex subpersonal calculation based on the weights of the various features in the prototype of *dog*. (The

details of this calculation differ a great deal depending on the particular version of the theory.)

The weight that a feature receives in the prototype is generally taken to be determined by two sorts of statistical facts, namely the *prevalence* of the property among *dogs* (so *barks*, *has a tail*, and so on, would receive high weights since most dogs have these features), and/or the *cue validity* of the property; that is, how likely it is that something with that feature is a dog. Thus even though, perhaps, most dogs don't wear collars, the probability of something being a dog if it wears a collar is high, so *wears a collar* might receive a significant weight in the prototype. The more highly weighted features an individual has, the more typical an exemplar of the kind it will be. Prototype theory thus places typicality effects first and foremost among the data it aims to explain.

Prototype theory has many adherents, and many well-motivated critics. For while we may often rely on statistically weighted features in categorization, particularly in rapid, perceptually based categorization, it seems that this cannot be the whole story. Imagine, for example, that you are presented with a raccoon. A perverse scientist then comes along and alters the creature, dyeing its fur so that it takes on the markings that are typical of a skunk, and even goes so far as to implant a sac of smelly liquid that the creature can use to spray smells when it is under stress. How would you categorize this creature? It now has all the typical features of a skunk, yet overwhelmingly, from elementary school on up, people say this is still a raccoon (Keil 1989). This finding has proven difficult for standard prototype theories to accommodate.

Furthermore, it seems increasingly clear that typicality ratings are not solely driven by statistical facts; crucially the causal status of features also matters. Imagine that two features are equally prevalent among members of a kind and have the same cue validity, but that one is understood as generally being the cause of the other. Suppose then that an instance of the kind has one feature but not the other. Since the relevant statistical facts are the same in both cases, prototype theory would seem to predict that typicality ratings of the individual would not be affected by *which* feature is lacking. However, individuals exhibiting the effect but not the cause are rated as less typical than those exhibiting the cause but not the effect (Ahn et al. 2000).

These results, along with many others, suggest that our ways of categorizing things, and reasoning about things in categories, involve a richly structured knowledge base that is responsive to causal-explanatory factors as well as statistical factors (e.g., Carey 1985, 2009; Gelman 2003; Gopnik and Meltzoff 1997; Keil 1989). As we will use the term here, this is the outlook typical of the so-called theory-theory of concepts.<sup>2</sup> Since theory-theory, so construed, posits that our concepts or criteria for

<sup>2</sup> Sometimes 'theory-theory' is reserved for the specific view that our psi-concepts, including young children's psi-concepts, are *very like* scientific theories. We follow many in the field by using the term more inclusively.

categorization and generalization are sensitive to causal-explanatory structure, there is not too much to be said beyond that about the *general* features of our concepts. Rather, it may be empirically useful to go on to consider concepts within each broad domain, e.g., natural kind concepts, artifact concepts, social concepts, mental state concepts, mathematical concepts, and so on and so forth.

For example, a view known as *psychological essentialism* seems to provide a great deal of insight into how our natural kind concepts are structured, indeed from a very young age (e.g., Gelman 2003; Leslie 2013). However, this view is very likely not applicable to artifact concepts or to mental state concepts, and certainly not to mathematical concepts. This sort of domain sensitivity should not be seen as a failing of theory-theory, but as an upshot of the complex and myriad ways we have of categorizing things, and of generalizing on the basis of those categories.

Probably the view that fits best with the mass of empirical material on concepts is a hybrid of theory-theory with some elements drawn from prototype theory. There are terms like "red" or "dog" which we can apply rapidly and without reliance on theory, at least in some circumstances. It is natural to think that this goes by way of subpersonal processing of sensory and perceptual information with subsequent comparison with paradigmatic or prototypical sensory and perceptual profiles—how red things look, how the varieties of dogs look, etc. The characterization of such prototypical profiles—in particular whether and to what extent they use prevalence and cue validity—is a complex piece of empirical psychological theorizing, yet to be completed.

## 7.2 THE GENERIC ENCODING HYPOTHESIS

One feature that the foregoing psychological theories of concepts have in common is that they all make some reference to properties that are possessed by members—*plural*—of the target category; they all involve focus on forms of *generalization* concerning the category and its properties. According to the classical theory, the relevant generalizations are universal generalizations; the prototype view treats them as probabilistic generalizations; while on the theory-theory, they are causal-explanatory general beliefs. These observations suggest a possible alternative route to studying the nature of our classificatory and inferential heuristics: we should look to the empirical investigation of our earliest and most fundamental types of generalization.

Suppose, for example, that it was possible to identify and describe our most basic and persistent way of forming general judgments about kinds or categories, our most basic and persistent way of moving from information concerning individual members of a category to judgments concerning the category in general. It would be quite surprising if this basic and persistent manner of generalization was not centrally connected with the heuristics for categorization and inference concerning kinds or categories. Thus a natural and conservative empirical hypothesis would be

that our conceptual heuristics in large part consist of such cognitively basic and persistent types of generalizations.

Recent interdisciplinary research suggests an intriguing possibility along these lines; namely that our basic way of generalizing information issues in *generic* generalizations, which are articulated in language via generic sentences such as “tigers have stripes,” “lions have manes,” and “mosquitoes carry malaria” (e.g., Gelman 2010; Hollander et al. 2002; Leslie 2007, 2008, 2012; Leslie and Gelman 2012; Leslie et al. 2011; Mannheim et al. 2011; Tardif et al. 2011).

Such generic sentences exhibit a puzzling truth-conditional profile, as a few familiar examples quickly illustrate. Consider, for example, “lions have manes”—this strikes most people as obviously true, yet only mature male lions have manes. There are perfectly normal lions (e.g., female lions) who lack manes, and yet the generic seems true. Further, there are *more* male lions than there are maned lions (since some males are immature or lack manes for genetic or environmental reasons), yet the generic “lions are male” is widely rejected, which suggests that contextual restriction is not the source of the phenomenon. Perhaps even more puzzling are generics such as “mosquitoes carry malaria,” which are accepted despite the fact that only about 1% of mosquitoes carry the virus. Yet generics such as “books are paperbacks” are robustly rejected, even though over 80% of books are paperbacks (for more discussion of generics, see Carlson and Pelletier 1995; Cohen 1996; Leslie 2007, 2008; for empirical investigation of people’s judgments of these sorts of generics and others, see Prasada et al. 2013).

Most importantly for our purposes here, generic generalizations are obviously not equivalent to universal generalizations, as is already confirmed by the truth of “lions have manes” and “mosquitoes carry malaria.” Obviously, even high-prevalence generics such as “tigers are striped” and “dogs have four legs” tolerate exceptions in a way that their universal counterparts do not. “All tigers are striped” is falsified by a single stripe-free albino tiger; similarly for “all dogs have four legs.” The corresponding generics are more robust, however. They are true in the face of such exceptions.

If generic generalizations constitute our most basic and persistent way of making general judgments about categories, this raises still another empirical challenge for the classical view. A proponent of the classical view would have to argue that the general information employed in our classificatory heuristics does not originate from our most basic and persistent way of forming general judgments. The information that we use to identify members of a category would not come by way of our basic and persistent means of forming general judgments about the category. This is not incoherent, but given the overwhelming absence of empirical evidence in favor of the classical view, it stands out as the positing of yet another defensive epicycle.

Likewise for what we have stigmatized as the quasi-supernaturalist picture of language use: if our classificatory heuristics commonly exploit generics, then our classifications are *not* invariably guided by a grasp of meanings understood as

determinants of universally necessary and sufficient conditions, i.e. conditions which would fix the extension of the classificatory term in question in every possible world.

Why should we think that generic generalizations are more cognitively basic and persistent than universal ones? Some of the relevant data come from the study of language acquisition. As noted, generics have a very complex truth conditional profile; providing an account of when generic sentences are true or false is a quite demanding task (see e.g., Carlson and Pelletier 1995; Cohen 1996; Leslie 2008). In contrast, it is very easy to provide an account of when universally quantified statements are true (“all Ks are F” is true iff the set of Ks is a subset of the set of Fs). In light of this, one would expect that universals would be easier for young children to acquire and process than generics; however, this is precisely the opposite of what we find. Generics are produced and understood by preschool-aged children, and the data collected to date suggest that these young children have a remarkably adult-like understanding of generics. For example, preschoolers who know that only “boy” lions have manes will accept “lions have manes” but reject “lions are boys”—despite implicitly understanding that there are at least as many “boy” lions as there are maned lions (Brandone et al. 2012; see also Gelman and Raman 2003; Gelman et al. 2002; Graham et al. 2011; for a summary of the available evidence on generic acquisition, see Leslie 2012).

Preschoolers are generally competent with the quantifier “all” *when it is applied to a specific set of individuals* (e.g., Barner et al. 2009). For example, if preschoolers are shown six crayons and asked, “Are all of *these* crayons in the box?” they are usually able to answer the question correctly. Most of the work on quantifier acquisition has focused on such situations; however, it should be clear that these sorts of limited, nonprojectable pseudogeneralizations are not the sort that are involved in conceptual heuristics. The question, then, is how young children fare with open-ended, category-wide universals—not “all of these crayons,” but “all crayons.” Several studies indicate that they have considerable difficulty processing universal quantifiers in such kind-wide generalizations. Most intriguing, though, is that when preschoolers are confronted with such kind-wide universals, they do not simply provide random, incorrect answers; instead *they treat the universals as though they were generics*. That is, preschool children not only consistently evaluate generics just as adults do, they also evaluate kind-wide universals as generics (Hollander et al. 2002; Leslie and Gelman 2012; Tardif et al. 2011; for a detailed review, see Leslie 2012). In addition to English-speaking children, such findings have also been documented among Mandarin Chinese- and Quechua-speaking children; similar results have also been found with other quantifiers (Brandone et al. 2015; Hollander et al. 2002; Mannheim et al. 2011; Tardif et al. 2011).

Importantly, these findings are just what one would expect on the hypothesis (Leslie 2007, 2008, 2012) that generics, unlike universals (and “some”- and “most”-quantified statements), articulate cognitively basic generalizations. If the cognitive system has a basic, default way of forming general, open-ended judgments, then

it may sometimes fall back on this means of generalizing when asked to process a more taxing and sophisticated generalization. This tendency will be most pronounced in young children, who would be expected to struggle with the more taxing generalizations. Not only do young children not struggle with generic generalizations, they substitute their understanding of the generic when asked to consider category-wide quantified generalizations.

If generics do in fact articulate cognitively basic, default generalizations, then one would expect that these effects might not be limited to young children. Adults might also be susceptible to the error of treating quantified statements as generics. Indeed, under a variety of circumstances, adults do show a robust tendency to accept universally quantified statements such as "All ducks lay eggs," despite knowing that male ducks do not lay eggs (where the tendency to accept the universal was *not* due to participants interpreting the universal as quantifying over only females, or over subkinds of ducks; Leslie et al. 2011; see also Meyer et al. 2011). This finding would be explicable if adults were not always evaluating the universal claim, but were instead, like preschoolers, sometimes substituting their evaluation of the corresponding generic.

Further confirming evidence for the persistence of the generic mode of generalization can be found in the study of adult reasoning errors. For example, Steven Sloman (1993, 1998) investigated adults' evaluations of arguments that involve the quantifier "all," finding that their evaluations did not conform to the logic of universal quantification. For example, his participants judged that arguments such as (A) are strictly stronger than arguments such as (B), despite judging that reptiles are indeed animals:

- (A) All animals use norepinephrine as a neurotransmitter; therefore all mammals use norepinephrine as a neurotransmitter  
 (B) All animals use norepinephrine as a neurotransmitter; therefore all reptiles use norepinephrine as a neurotransmitter

This pattern of judgment is simply mistaken given the logic of the universal quantifier; however, if we replace the universals in the arguments with generics, then the judgments of the participants would be very reasonable. Since generics tolerate exceptions, the claim "Animals use norepinephrine as a neurotransmitter" can be true even if some animals are exceptions to the claim. If one also judges that reptiles may not generically use norepinephrine while mammals probably do, then argument (A) is indeed stronger than argument (B). Hence these results are as one might expect if adults have a tendency to evaluate universals as generics.

Note that adults also judge that universals such as "All ravens are black" are more likely to be true than universals such as "All young jungle ravens are black," despite understanding that the latter are a subset of the former (Jönsson and Hampton 2006). Again, this is incoherent if one is really dealing with universally quantified

statements; however, if one were instead evaluating these universals as generics, this would be a reasonable judgment, since for all one knows young jungle ravens may be exceptions to the generic "ravens are black." These results are thus naturally read as lending support to the hypothesis that adults are treating these universals as generics. As a further piece of converging evidence from another experimental paradigm, it has been found that both preschoolers and adults *recall* previously presented quantified statements as generics (Leslie and Gelman 2012).

The hypothesis that generics, unlike quantified statements, articulate cognitively basic and persistent generalizations thus has a fair amount of empirical support at this time. As a final observation in favor of the hypothesis, we might note that quantified statements require a phonologically articulated element, namely the quantifier itself. That is, we say "All tigers are striped" or "Most tigers are striped"; however, in the case of the generic, there is no corresponding articulated element (e.g., "Gen tigers are striped"). This is not an isolated fact about English; rather it would appear that few, if any, natural languages have a dedicated, articulated generic operator (Carlson and Pelletier 1995; Dahl 1985).

Here is an explanation for this otherwise puzzling fact: if one wishes to interact efficiently with a system, and the system has a basic, default way of proceeding or performing a task, then one need only issue an explicit instruction to the system if one wishes it to *deviate* from this default way of proceeding. To convey the idea in more intuitive terms, if one is dealing with a child who, say, by default does not pick up her toys, one only needs to say something if one wishes the child to *deviate* from her default and actually pick up her toys. If one does not wish the child to pick up her toys on a given occasion, it would be a waste of breath to say, "Don't pick up your toys!" since this is what will happen even if one remains silent. Thus, quantifiers may be articulated in language because one needs to *tell* the cognitive system, as it were, to deviate from its default, generic mode of generalizing, and instead generalize in the universal manner or the existential manner and so on and so forth (for more details, see Leslie 2008, 2012). Generics, by virtue of expressing the basic or default mode of generalization, require no such phonological marking. They represent the unmarked case.

More can be said about the centrality of generic generalization (see Leslie, forthcoming). Suffice to say, the best available theoretical construal of the data supports the view that our first inferences to generalizations about kinds produce generics, and that the generic form remains the default mode of generalization, even in adult life. Second, as already noted, on both the theory-theory and prototype theories, the heuristics which guide the application of our terms are general in form; indeed they must be so if they are to be available for use from one occasion to another.

Those two considerations lead us to propose a new, empirically motivated and philosophically consequential, amendment to both the theory-theory and the prototype theory, namely the *generic encoding hypothesis*: the heuristics which typically

guide our use of terms by exploiting prevalence, cue-validity,<sup>3</sup> and causal explanatory structure<sup>4</sup> are properly formulated in generic terms.<sup>5</sup>

### 7.3 PSI-CONCEPTS ARE NOT MEANINGS

We are now in a position to see why the cluster of generic heuristics a speaker associates with a term should not be taken as specifying the semantic meaning of that term in the speaker's language.

For one thing, the exact heuristics employed may vary from speaker to speaker, each of whom may speak the same language *L* because of their convergence on a pattern of conventional use, a pattern corralled by deference to correction by distinguished others, i.e. parents, teachers, and those in the know. They may then come to mean the same by a given term of *L*, even though they employ different heuristics. The same for a single speaker over her lifetime: without changing the meaning of a term *F* she may simply learn more about *F*s and thereby come to have better recognitional heuristics when it comes to classifying *F*s; even so, she may continue to use *F* with its *L*-meaning, i.e. continue her earlier pattern of conventional use of *F*.

Second, even when an inference from

*X* is *F* and *X* is a *G*

to

*X* is an *FG*

is *L*-valid, i.e. truth-preserving in virtue of the meanings of the terms of *L*, it may not be the case that the common heuristic for an *FG* among *L*-speakers is a union of their individual heuristics for *F* and *G*. This can be so, even though *FG* is a term whose extension is just the intersection of the extensions of *F* and *G*. For *FG*s may lack one or another of the classificatory heuristic features had by either *F*s or *G*s.

Here is a nice example adapted from Jerry Fodor and Ernie Lepore (1996). If some specific thing is both a pet and a fish then it is a pet fish. But the prototype of a pet fish, or as we would put it, the most salient generic pet fish, is a goldfish, which is not well modeled by the union of the features of generic fish and generic pets.

<sup>3</sup> Here we have in mind such heuristics as "Collared animals are dogs."

<sup>4</sup> "An animal's outer appearance is due to its 'interior' properties."

<sup>5</sup> It is, perhaps, rather telling that psychological experiments concerned with concepts almost invariably use the generic form to articulate conceptual knowledge.

There is a third way in which concepts, understood as clusters of heuristics associated with a term, may come apart from meanings. Two people may associate the very same heuristics with *F*, and yet mean different things by *F* in the sense that their dispositions to use *F* in various speech acts to contribute to communicative ends, i.e. to things being asserted, commanded, inquired after, etc., can differ. For their patterns of deference to the usages of parents, teachers and experts can significantly differ. Ms. Self-Sufficient may not be into that sort of thing at all, while Mr. Go-Along-to-Get-Along may be totally open to correction by others.

Suppose that both Ms. S's and Mr. G's heuristics for applying the term "dog" are exhausted by the following down-and-dirty criteria:

Is it an animal?

Does it have one of the characteristic looks, smells, coat textures, etc., of one of the familiar kinds of animals we call 'dogs'?

Is it the offspring of an animal with one of the characteristic looks, smells, coat textures, etc., of one of those familiar kinds of animals we call 'dogs'?

As things actually go in suburban environments, these three criteria may be good heuristics for collecting together observed instances of the kind dog, the kind we now know to be the species *Canis familiaris*. However, this is due to the contingent fact that the canines around us are almost all of them from that species. The dogs have a sister taxon *Cana lupus* which includes all and only the wolves. Wolves are not dogs, but some wolves look very like German shepherds, some look very like huskies, some look very like malamutes, and some look very like the new dog breed Tamaskans, which were deliberately bred from German shepherds, malamutes, and huskies to resemble wolves.

However, wolves avoid human communities and so typically don't roam in suburban neighborhoods, like those frequented by Ms. S and Mr. G. So, both Ms. S and Mr. G invariably recognize the dogs around them as dogs, and since they do not encounter wolves they do not in fact count any wolves as dogs, though both would be likely to do this, since they lack distinguishing criteria, such as that dogs bark while wolves howl. What makes Mr. G's term "dog" refer to the dogs rather than to a group which includes both the dogs and the wolves is his pattern of deference to distinguished use—in this case to that system of biological classification which tells us that the dogs form a species *Canis familiaris*, a species which excludes the wolves. However, Ms. S, who has no such disposition to defer, has nothing that would make her term "dog" have such an extension. Given just her criteria or ways of telling, her term "dog" has both dogs and doglike wolves in its extension. In this respect, though their heuristics are the same, Ms. S and Mr. G speak a different language in the strict sense of a language understood as a potentially infinite list of pairings of sentences and (extension-determining) meanings. For Ms. S's term "dog" arguably has both dogs and wolves in its extension.

## 7.4 PHILOSOPHICAL CONSEQUENCES

Though we “know the meanings of our terms” this is typically not *by way of* grasping meanings. Semantic meaning properly figures in the individuation of a given language *L*, in the characterization of analytic validity in *L*, and collaterally in the individuation of “concepts” semantically conceived, i.e. as pairings of (subsential) terms with their meanings. It has at most a minor role in a general explanatory account of how we use terms. A very small percentage of simple terms are like “vixen,” “copse,” etc., where speakers *are* in a position to articulate the terms’ semantic meaning.

As one of us argued long ago (Johnston 1988), although questions of syntax are of great interest, the pure theory of semantic meaning has very little to it. Semantic meanings are functional indexes, suited to play certain quasi-model theoretic roles in the specifications of truth conditions and of analytic validity. To think that it is grasp of semantic meaning that quite generally guides our use of terms is a level-confusion. The truths about the semantic meanings of our terms are epiphenomena generated by our conventionally constrained use of those terms. And this use is guided by our generic, prototypical, and partly causal-theoretical criteria for applying and withholding the terms.

So what?

Well, to begin on exploring the philosophical consequences of this view, recall verificationism about meaning: the view that the meaning of a sentence is to be given by its methods of verification and/or falsification, which in turn decompose into the heuristics which guide our application of the subsential expressions which make up the sentence. In the light of the foregoing, verificationism appears to be another form of the level-confusion we have been emphasizing. Methods of verification and falsification—heuristics or criteria—are one thing, meanings another.

When we apply any predicate *F*, we are using criteria or ways of telling—i.e. methods of fairly direct verifying and falsifying—whether such and so is an *F*. If our respective ways of telling quite generally coincide in their deliverances, and if dispositions to defer also coincide, we may come to speak the same language, i.e. to be disposed to use the sentences of some language with the same speech act potential, namely to assert, command, inquire after, etc., the same things. And yet, thanks to the language’s rules of semantic composition, there may be meaningful sentences of the language, such as unqualified universal generalizations, which have no method of direct verification, along with other sentences, such as negative existentials, which have no method of direct falsification. There may be meaningful sentences of the language about events in the past, all of whose traces have washed away. Likewise for sentences about events in “pocket universes” that are forever outside our light cone. We could have a well-confirmed theory that tells us there are such events, while leaving their specific character open. The unverifiable and unfalsifiable sentences representing the *specific* characters of such events nonetheless will have a meaning

compositionally conferred upon them. As a result, some of the things that speakers can do with the language is ask questions which have no accessible answers. The corresponding possible answers thus have meanings that transcend what can be settled by speakers’ methods of verification and/or falsification for the sentences that express them.

Recall that Michael Dummett held that a theory of meaning for a language was a theoretical representation of a practical ability by way of a statement of what, knowledge of which, would suffice to be able to use the language. He then famously argued that a theory of meaning should take the form of a statement of conditions of verification and not an assignment of potentially evidence-transcendent truth conditions to the sentences of the language. For, he reasoned, the latter would overdescribe the practical ability we do in fact have, precisely because our criteria of use underdetermine such truth-conditions.

In the light of the forgoing, that argument looks like the familiar mixing of levels, the level of meaning and the level of what guides concept application. One thing we can do with our language is ask questions that have epistemically inaccessible answers. This ability is conferred on us not by our actually being guided by evidence-transcendent heuristics (whatever they would be), but by being able to produce speech acts with evidence-transcendent truth or satisfaction conditions, thanks to the semantically compositional character of the language we speak.

Recall that Jerry Fodor was keen to attribute verificationism to cognitive psychologists offering empirical theories of concepts. As against Fodor’s useful provocations, we have argued here that the proper theoretical placement of that psychological work shows precisely why verificationism is mistaken and why the relevant cognitive psychologists were *not* in fact verificationists. Fodor’s wholesale rejection of psychological theorizing about concepts, so far as we can see, consisted of making just the right points about concepts in the philosophical sense and then supposing that psychologists were theorizing about that. Again, the levels have to be confused to get the argument going.

Despite the storytelling that presents prototype theory as an improvement on the classical theory, which in its turn *seems* like a meaning theory, the *psi*-concepts are not an empirical account of the *phi*-concepts, nor were they meant to be.

A related line of thought emerging from our reflections suggests that names could have meanings, over and above their denotations, even though Saul Kripke (1980) was entirely right that the meaning of a name is not given by our ways of identifying the bearer of the name. Setting aside Kripke’s modal argument, the success of which turns on the question of whether our heuristics are properly understood as tied to how things go in actuality, Kripke’s other arguments—that our heuristics differ, that they are often impoverished, and that they are ill-suited in the specification of universal necessary and sufficient conditions for being the bearer of the name—are from the present point of view best understood as perfectly correct antiverificationist points. As Kripke himself was careful to note, those arguments



do not in themselves take us all the way to "Millianism" about names, i.e. the thesis that names have no meaning over and above their denotation. They simply rule out identifying the meaning of a name in our language with the heuristics we use to apply it. In our terms, Kripke highlighted the difference between psi-concepts and phi-concepts in the case of proper names.

## 7.5 THE NORMATIVITY OF MEANING

Another upshot of distinguishing psi-concepts and phi-concepts has to do with the so-called normativity of meaning, also brought into sharp relief by Kripke (1982). To be sure, semantic meaning is normative in that there are correct and incorrect uses of terms. But this is so only *relative* to some specific language. "To table" in American English means to remove from the main line of discussion. In British English, it means the opposite: to introduce into the main line of discussion.<sup>6</sup> (Apparently, the mutual failure to understand this led to a temporary rupture between Churchill and Roosevelt during an important campaign of the Second World War.) Clearly the norms for correct use have to be stated relative to a language; they are conditional norms, such as

If you aim to use 'to table' as an expression of American English, e.g. in a conversation with speakers who only have American English, then you should use it to mean: to remove from the main line of discussion.

If you aim to use 'to table' as an expression of British English, e.g. in a conversation with speakers who only have British English, you should use it to mean: to introduce into the main line of discussion.

There is accordingly nothing puzzling about the so-called normativity of meaning; it is just a case of the normative principle that tells one to use a means appropriate to one's end. The intuition that the normativity of meaning is somehow more than this, indeed somehow *categorical* in nature, can be traced, we suggest, to the relative paucity of homophones with differing meanings across different languages. Because of this contingent fact, the always correct antecedent reference to a language is almost always harmlessly left out. The "normative force" you feel is the internalization of conventional pressure towards local conformity of use; it is not the supernatural influence of a meaning rule on your psychological life. There is no general semantic conscience to guide us. ("Vixen," "copse," and the like notwithstanding.) What we sometimes feel to be guiding our use are our heuristics, not antecedent *graspings of meaning*.

As far as following a linguistic rule goes, the psychology of language use is the place to look. It implies that there is no guarantee of singularity or of determinacy

<sup>6</sup> Thanks to Christopher Peacocke for the example.

when it comes to the rule to which our use of a term conforms. More generally, there is no guarantee that there is a unique language in the abstract sense that we are related to by the actual language relation.

This is not "meaning skepticism" in any worrying sense, precisely because meaning is not the source of use. Meaning simply figures in the systematic *registration* of that use for formal semantic purposes, i.e. individuating a language L, specifying conditions of truth-in-L, and characterizing analytic validity-in-L.

## 7.6 THE GENERIC ENCODING HYPOTHESIS AND THE METHOD OF CASES

As late as 2010, Michael Dummett can be found defending the view that the proper method of philosophy is the analysis or articulation of the conditions of application of our concepts. Several of our own colleagues, especially among those working in ethics, treat their topic as nontrivial conceptual truth and their method as the codification of conceptually clearheaded reactions to cases.

The background thought may be developed as follows. As masters of concepts we have at least an implicit grasp of their application conditions; this tacit knowledge of when they apply and when they should be withheld can be manifested equally well in real and imaginary cases. This must be so, since the master of a concept is antecedently armed with a capacity to tell whether or not to apply the concept, however reality might turn out to be (perhaps within certain limits of normality). Here then is a method for articulating our tacit knowledge of the application conditions of our concepts.

In the best scenario, the method delivers a "conceptual analysis"; that is, an account of a special sort of universally necessary and sufficient condition or set of conditions for the application of the relevant concept, namely a universally necessary and sufficient condition or set of conditions that could be recognized as correct simply on the basis of a certain sort of ideal reflection on our tacit understanding of when to apply and when to withhold the concept in question.

Therefore, the relevant verdicts and the resultant analysis can be delivered from the armchair, i.e. without any significant empirical investigation; so it is sometimes said that the relevant analyses could be known a priori; roughly, in a condition approximating to blissful ignorance of the empirical facts.

There were a few promising victories for this kind of method, but they were skirmishes rather than major battles. The analysis of the concept of knowledge was, at least for a while, considered a paradigm of this kind of investigation, one which neatly exemplified how "the method of cases" could lead us to an analysis of a concept. "Intuitions"—that is judgments—as to whether the case at hand was, or was not, a case of knowledge were collected by visiting real and imaginary cases alike, and then those intuitions were brought into some sort of reflective equilibrium that bore on the question of the universally necessary and sufficient conditions for someone's

knowing some arbitrary proposition. Imagined cases were naturally treated as on a par with real cases; for if we are interested in articulating our tacit understanding of the application conditions of our concept it would be odd to restrict our evidence base to the adventitious experiments of stepmotherly nature, when we could also avail ourselves of the full range of ingeniously designed thought experiments. Wouldn't that be like only considering the moves that have been made in actual chess games, rather than the full range of moves that *could have been* made? As in chess, so with our concepts: imagination is a reliable guide to what could happen. It thereby provides us with cases that are just as helpful as the actual cases so far as rendering explicit our implicit understanding of the application conditions of our concept; as it might be, the concept of a mate in four or the concept of knowledge.

The ideology behind the method of cases thus offers to explain how the imagination can have a probative status, how it could have a kind of evidential significance which mere fancy could not. The imagination's philosophically interesting function is to generate a wider than actual range of cases, across which our conceptual competence can express itself.

For a good while, this method looked attractive when it came to the concept of personal identity. The Anglophone philosophy of personal identity emerged as a going concern in the 1960s thanks to the work of such philosophers as David Wiggins, Bernard Williams, Sydney Shoemaker, John Perry, Derek Parfit, and others inspired by them. These philosophers worked explicitly within the idiom of analytic philosophy and supposed that the real task of the philosophy of personal identity was to illuminate our *concept* of personal survival by means of organizing our intuitions about survival or continued existence, intuitions gleaned from a wide range of real and imaginary cases.

The fact that the target was a *concept* made the method of cases look like a viable approach in the case of personal identity. We are highly competent with the concept of personal identity; we have applied it successfully in a wide range of cases throughout human history, and in the common run of cases we appear to have a mass of accumulated knowledge of who is, and *was*, whom. So we must have at least an implicit grasp of the application conditions of the concept of personal identity, and this tacit knowledge of the concept's application conditions can be manifested equally well in real and imaginary cases.

Thus in the case of the concept of personal identity the dominant method in analytic philosophy was then to collect intuitions about real and imaginary cases of personal survival and ceasing to be, and then bring those intuitions into some sort of reflective equilibrium that bore on the question of the necessary and sufficient conditions for an arbitrary person's survival. The result would be the filling in of the details of the relation *R* in an a priori (and necessary) biconditional of this form:

*x*, considered at *t*, is numerically the same person as *y*, considered at *t\**, if and only if *xRy*.

A specification of *R* would entail a specification of the "identity" or, better, "persistence" conditions of persons, i.e. a specification of what changes they could and could not survive. Thus arose the old analytical question: Is *R* a matter of *x* and *y* having the same body, or being the same organism, or having the same consciousness, or having the same mind (however that mind might be embodied), or having the same separable immaterial soul?

There are many worries that have been and can be raised against this whole approach to the question of personal identity,<sup>7</sup> but we are now in a position to present a dilemma that threatens the whole approach: *is the target concept a phi-concept or a psi-concept?*

If the concept of personal identity is a psi-concept, shouldn't we admit that it is the techniques of cognitive psychology, some of whose interesting results have been already described, which provide the best way to get at our concept, i.e. the cluster of heuristics which guide us in the use of our terms?

On the other hand, if it is a phi-concept or meaning, then absent a quasi-supernatural picture which depicts our use of terms as quite generally the upshot of grasp of meaning, why should we suppose we are in a position to articulate those meanings just from armchair reflection? We do not know how to use our terms by way of knowing a theory of meaning for our language. It is not just that, as Hilary Putnam once famously said, "Meanings are not in the head." Meanings are not anywhere at all; they are theoretical posits for the purpose of individuating languages, and defining language-relative truth conditions and validity. Meanings register usage; they do not generate it.

In response to this dilemma, many philosophers today would abandon the second horn. After all, there has been a widespread loss of hope that any true biconditional of the above form could be analytic, so that those who denied it would either be failing to understand their language or speaking another language.

But then, whence the conviction that it is philosophers who are best placed to articulate the relevant heuristics which have provided us with knowledge of who is, and *was*, whom in the common run of cases? Whence the conviction that these heuristics, taken individually or collectively, apply a priori? Couldn't they all be empirically falsified, so that the articulation of our psi-concept might amount to simply a folk-theory of personal persistence, one open to empirical refutation?

The best answer respects the good thought that we cannot always have been totally wrong about something, on pain of losing the something to have been wrong about. For how then would we have focused on a topic to be totally wrong about? (Setting aside the infrequently used method of bare perceptual demonstration of an item as "this" or "that" without having any controlling conception in play.)

However, the good thought is a holistic one, with minimal constraining power. The good thought leaves room for massive error.

<sup>7</sup> For some others, see Wilkes 1988 along with Johnston 1987 and 2010, 44–47.

More particularly, the generic encoding hypothesis allows for the possibility that we could be right enough about a topic without knowing any very interesting universal truths about it, but only useful generics involving it. This suggests that the urge to articulate universal necessary and sufficient conditions concerning the topic by way of articulating our criteria or heuristics has no general theoretical justification, and may just be misguided.

That this is not just an abstract possibility is suggested by the case of personal identity itself. It may be that our heuristics for tracking a persisting person involve the generic criterion to the effect that persons survive if their individual minds continue on, and the generic criterion to the effect that persons survive if their bodies are kept alive and functioning. But it may also be that when these two heuristics point in different directions we should shrug our shoulders, and are inclined to do so, unless operated on by some misleading presentational feature of the case. So many of the imagined cases in the philosophical literature, such as teletransportation and brain transplantation, may simply be cases in which our psi-concept idles thanks to the generic character of the heuristics involved. The same may hold for some actual cases, such as the persistent vegetative condition. The result is that there is simply no analysis of the concept of personal identity, as traditionally understood.

So also with other concepts of philosophical significance. For all Gettierology showed, "Knowledge is true justified belief" and "Justified true belief is knowledge" may both be true generics. There is no general reason, arising from the nature of grasping meaning, to suppose that there is an a priori universal statement of necessary and sufficient conditions somewhere in the offing. The generic encoding hypothesis makes the ambition of traditional philosophical analysis seem quaint for most concepts.<sup>8</sup>

We take these reflections to reinforce the idea that the proper philosophical method, here as elsewhere, is not to limit oneself to the impoverished realm of conceptual or a priori knowledge, knowledge somehow deriving from, or embedded in, our competence with the meaningful terms of our language. The proper method is to use *all* one knows and all one can find out, in the most ingenious ways one can. Philosophy is not, and certainly not exclusively, the analysis of concepts. Philosophy is integrative theoretical vision combined with argumentative ingenuity, deployed at a fairly abstract level. Who could fail to love that, at least when it is done well?

On this view, philosophy has *no* special "resource" in Michael Dummett's intended sense. But so far from marginalizing philosophy, this liberates it. The upshot is merely a clarification of what was always anyway our obligation as philosophers: we must learn a lot more science than the analysts of yore deemed relevant.

We need to get out of the armchair and again start looking into things.

<sup>8</sup> Similarly, "Generically speaking, you do not make a hirsute man bald by removing a hair" seems a good heuristic guiding our use of 'bald'. But . . . well, you know how that goes. A question: how much of how that goes turns on the confusion between psi-concepts and phi-concepts?

## 7.7 THE END OF MEANING

Once psi-concepts are separated off from semantic meaning, a question remains as to the interest of semantic meaning. Suppose that as a matter of fact, the actual examples of analytic validity are rather few and far between, and that conventionally constrained use is not what it was (if it was ever what it was!) in the sense of producing a stable conformity of use over significant periods of time. Then the notion of a population sharing a language in the semantically tight sense, though perfectly coherent, may find little de facto application.<sup>9</sup> We would then look back on the theory of meaning as a massive overapplication of the tools of model theory, without much in the way of results to show for it.

Less pessimistic but still important is the claim we have argued for here. Cognitive psychology undermines the metaphysics of meaning, the idea that use is *guided* by grasp of meaning. In doing so, it also lays to rest the analytic conception of philosophy with its characteristic target of analyses facilitated by the method of cases. This result stands even if there remains an analytic/synthetic distinction, along with some inferences that are valid in virtue of meaning.

## REFERENCES

- Ahn, W., Kim, N. S., Lassaline, M. E., and Dennis, M. J. (2000). Causal status as a determinant of feature centrality. *Cognitive Psychology* 41: 361–416.
- Armstrong, S. L., Gleitman, L. R., and Gleitman, H. (1983). What some concepts might not be. *Cognition* 13: 263–308.
- Barner, D., Chow, K., and Yang, S. (2009). Finding one's meaning: A test of the relation between quantifiers and integers in language development. *Cognitive Psychology* 58: 195–219.
- Braddon-Mitchell, D., and Nola, R. (2009). *Conceptual Analysis and Philosophical Naturalism*. Cambridge, MA: MIT Press.
- Brandone, A., Cimpian, A., Leslie, S. J., and Gelman, S. A. (2012). Do lions have manes? For children, generics are about kinds, not quantities. *Child Development* 83 (2): 423–33.
- Brandone, A., Gelman, S. A., and Hedglen, J. (2015). Young children's intuitions about the truth conditions and implications of novel generics. *Cognitive Science* 29: 711–38.
- Carey, S. (1985). *Conceptual Change in Childhood*. Cambridge, MA: MIT Press.
- Carey, S. (2009). *The Origin of Concepts*. New York: Oxford University Press.
- Carlson, G. N., and Pelletier, F. J. (1995). *The Generic Book*. Chicago: Chicago University Press.
- Cimpian, A., Brandone, A. C., and Gelman, S. A. (2010). Generic statements require little evidence for acceptance but have powerful implications. *Cognitive Science* 34 (8): 1452–82.
- Cimpian, A., Gelman, S. A., and Brandone, A. C. (2010). Theory-based considerations influence the interpretation of generic sentences. *Language and Cognitive Processes* 25 (2): 261–76.
- Cohen, A. (1996). *Think Generic: The Meaning and Use of Generic Sentences*. PhD dissertation, Carnegie Mellon University.

<sup>9</sup> This may be the view that Donald Davidson (1986) expresses in his important but elusive essay "A Nice Derangement of Epitaphs" by provocatively suggesting that there are no languages. Whether the contemplation of malapropisms is the most direct route to this result is another question.