AN ORIGINALIST THEORY OF CONCEPTS

R. M. Sainsbury and Michael Tye[1]

Philosophy Department, University of Texas at Austin

The main idea in this paper is that atomic concepts are to be individuated by their historical origins. Nonatomic concepts and thoughts are individuated by their constituent concepts and how they are structured. This originalist position contrasts with standard views, according to which concepts are to be individuated semantically or epistemically. Unlike Millianism, originalism is able to take Fregean data at face value. It also gives natural descriptions of cases, like Mates cases and examples involving demonstratives, that are problematic for Fregean views. Individuating concepts in a way that does not depend on semantic or epistemic matters makes concepts available to provide non-circular and simple explanations of puzzles concerning thought.

Concepts are vehicles of representation, tools for thinking. They are individuated historically, and can be combined into structures we call *thoughts*. Thoughts can be evaluated as true or false. Typically, concepts have reference: for example, nominative concepts typically refer to objects, predicative concepts to properties, and so, indirectly, to objects possessing those properties. Some concepts fail to refer, but this does not prevent them having a role in thought. Distinct thoughts, even if they are referentially isomorphic, can play different cognitive roles. This enables us to make room for Fregean data, for example that the thought (in our sense) that Hesperus is visible is distinct from the thought that Phosphorus is visible. The concept <u>Hesperus</u> and the concept <u>Phosphorus</u> were introduced on distinct occasions, one at dusk, the other at dawn, so they are distinct. For similar reasons, the thought that Pegasus is a horse is distinct from the thought that Vulcan is a horse, even though there is no difference at the level of reference.

---

[1] This is fully collaborative work, and no significance attaches to the (alphabetical) order of the authors' names.

1. *Origins*

Atomic concepts come into existence (and perhaps go out of existence); they are non-eternal continuants. For example, the concept <u>quark</u> did not exist in 1900 but it exists now. Every theorist must make room for such plain matters of historical fact.

A given individual may use a concept for the first time in one of two ways. It may be the first time the concept has ever been used. In this case, the use constitutes the introduction of the concept, and we call such a use an *originating* use. Alternatively, the concept may already be in existence, and the individual's first use of it constitutes his becoming a member of an existing concept-using community. In this case, the use is not an originating use of the concept.

If we speak of the originating use of a specific concept, <u>C</u>, we seem to presuppose the identity of <u>C</u> itself. Our first task is to characterize an originating use in a way that does not presuppose prior identification of the concept thereby originated. This will pave the way for the key positive thesis of originalism: concepts are individuated by their originating use.

We begin by thinking about non-originating uses. There are two sufficient conditions for a use to be non-originating:

1. The use involves deference to other uses, by the same subject or other subjects.
2. The use involves informational accumulation from other uses, by the same subject or other subjects.

When Gell-Mann originated the concept <u>quark</u>, he was in no way deferring to other uses of the concept, whether by himself or by other scientists. By contrast, our current uses of the concept <u>quark</u> involve deference: we aim to conform our usage to our previous usage, and to the usage of those in our conceptual community, especially to the usage of the scientifically informed. This is typical of many non-originating uses. The deference takes the form of intending to use the concept as it has been used by oneself or others on previous occasions. Using it in the same way does not require one to use it to think the same thoughts: one can change one's mind or disagree with others. But this kind of

change of mind or disagreement requires agreement in concepts. Shifting concepts would be a kind of equivocation, rather than a genuine change of mind or disagreement. In a non-originating use, a suitable kind of conformity with an existing practice is required, and this must exercise some normative force. For example, conformity is intended, so failure to conform is failure to act as one intended, and so involves a mistake. One characteristic of originating uses is the absence of any such conformist requirement.

In many non-originating uses, information is brought forward from other uses; it is "accumulated". This is manifest in adult uses: to learn about quarks is to make further use of the concept quark in the course of adding to the information the concept is used to register. It is also a feature of infant cognition. Studies suggest that infants respond to, for a example, a red block, not merely with separate thoughts, ones we might verbalize as "That's red" and "That's a block", but also with conjoined thoughts, in which the information is accumulated: both the information *is red* and the information *is a block* are subsumed under a single concept.

Originating uses are ones not governed by any conformist norm (for example, not involving deferential intentions) and not involving any informational accumulation. They are of various kinds. Explicit intentional introductions of concepts, like Gell-Mann's introduction of the concept quark, are easy enough to describe and understand. However, they involve intentional actions, and having intentions of the kind in question requires using concepts. Hence not all concepts can come into existence through explicit intentional introductions.[2]

Originating uses that are not explicit intentional introductions arise in normal behavioral and neurological development. The cognitive architecture of many species, including our own, ensures that concepts arise in young organisms as part of normal development, prior to language use. Developmental psychologists study when babies first come to possess concepts, and what the concepts are concepts for. Originating uses of these concepts occur through the operation of subpersonal processes. They

---

[2] We assume that we are not born with concepts, though we are of course born with mechanisms which, as they mature and interact with the environment, generate concepts.

involve nonconceptual contact between the mind and objects of thought. That there are nonconceptual ways in which the mind can make contact with other things has been argued for theoretically (Tye 2005, forthcoming) and hypothesized in the course of empirical work (Pylyshyn 2007). This is fortunate, for otherwise it would be hard to explain how concepts could arise in a being lacking concepts (this was Fodor's puzzle, leading him to an extravagant innateism).

In 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. An infant might originate a concept for cats. But we believe that as the child becomes a member of his conceptual community, this concept will be supplanted by one or more public concepts, for example the concept cat. Linguistic immersion is also conceptual immersion. In learning and coming to use the public language in our community, we thereby acquire and come to use the concepts these words express. This is the source of the majority of the concepts we use.

Here is a key originalist claim:

O1   For every concept, there is just one originating use.

Does this contrast with the way in which the everyday notion of a concept is used? For example, we might say that Leibniz and Newton both independently devised the concept integration (as used in the calculus). According to (O1), by contrast, if both Newton's and Leibniz's uses were originating uses, two concepts were originated. We suspect that ordinary usage does not reliably mark distinctions which we regard as important. A *conception* of water is a body of information concerning water. There is no such thing as *the* concept *of* water (various distinct concepts, like the concept H_20 and the concept stuff that falls as rain, have water as their referent, and so are concepts *of* water). By contrast, there is such a thing as the unique concept water. Newton and Leibniz may both have formed concepts *of* integration, and may even have had similar or the same *conception* of integration. It does not follow, and on our account it could not be true, that they both formed the concept integration.

2. Individuation by origin

After an originating use of a concept, there are (zero or more) "descendant" uses: uses that count as uses of that very same concept thanks to standing in the ancestral of a deference-involving relation R to the originating use. We first came into contact with the concept quark at a lecture by Gell-Mann. When we first used it, we were trying to use it as he used it in the lecture. We have maintained those efforts. Maybe we have now forgotten that early learning experience, but in trying today to use the concept as we used it yesterday we are maintaining a link of dependence with our first use; and so, indirectly, with Gell-Mann's originating use.

Originalism answers the question: what are the necessary and sufficient conditions for the concept C1 to be the same concept as the concept C2? According to originalism, every concept has exactly one originating use[3], and every originating use of a concept is an originating use of just one concept. Hence we can offer the following necessary and sufficient condition for concepts to be the same:

(O)   Concept C1 = concept C2 iff the originating use of C1 = the originating use of C2.

In this picture, each use U of a concept is a use of the unique concept that lies at the origin of the R-linked chain of uses to which U belongs. From the point of view of metaphysics, all facts of identity and difference among concepts are settled. We do not, of course, suggest that this is how we in practice distinguish concepts: our originalism fixes the metaphysics of concepts, not their epistemology.

Individuation by origin does not reveal the nature of concepts, nor does it reveal what makes a specific concept the concept it is. We take it that one can begin to explain the general nature of concepts in terms of thoughts, of which concepts are constituents, as we did at the start of this paper (and about which we say more in §4 below). As for

---

[3] For all we know apriori, there is an infinite human past, and so, for many concepts, no moment at which they were introduced. It is an indisputable aposteriori fact that our concepts are not like this (thanks to Brian Cutter for the observation). Perhaps a claim both fully general and apriori would be that concepts are individuated by their unique histories.

the more specific question, for example, what makes the concept <u>Hesperus</u> that very concept, rather than any other, we hold that this has no reductive answer. A correct answer is that it's the unique concept using which enables one to think of Hesperus as Hesperus; but this, harmlessly because inevitably, re-uses the concept <u>Hesperus</u> in the explanation.

3. Contents

There are many different kinds of atomic concepts: nominative concepts like the concept <u>Hesperus</u>, predicative concepts like the concept <u>barks</u>, concepts for logical operations like the concept <u>and</u>, concepts associated with intensifiers like the concept <u>very</u>, and so on. The relevant taxonomy mirrors the appropriate taxonomy for public language expressions, expressions for concepts. Our official view is that concepts should have the semantic properties that are attributed, by the best semantic theory, to the words that express them. Natural language semantics is an ongoing project, so we cannot help ourselves to a finished semantic theory. However, for illustrative purposes we will adopt a very simple-minded style of semantics. Our aim is to show that one does not need sophisticated semantics to explain Fregean data and related puzzles. Indeed, the main thrust of our originalism is to show that most such puzzles are to be explained not in terms of semantic differences, but in terms of conceptual differences. We highlight this point by adopting very coarse-grained contents for concepts, and by allowing that there are concepts that can be used successfully in thought while lacking content altogether.

In this illustrative vein, we will presume that the contents of contentful atomic nominative concepts, like the concept <u>Hesperus</u>, are the individual objects to which the concepts refer, like the planet Venus; and the contents of contentful atomic predicative concepts, like the concept <u>being happy</u>, are the properties which the concepts introduce, like the property of being happy. Nominative concepts may lack any content, like the concept <u>Pegasus</u>. Likewise for predicative concepts: one may hold a theory of properties according to which there is no such property as *being phlogiston*, but this would be no barrier to the existence of the concept <u>phlogiston</u>.

The concept Hesperus has the same content as the concept Phosphorus, since they refer to the same thing. However, the concepts are distinct, having different origins. This explains how it can be one thing to think that Hesperus is visible and another to think that Phosphorus is visible, and how it can be one thing to think that Pegasus is a horse and another to think that Vulcan is a horse.

What facts about the use of a concept ensure that it has the content it has, or that it has no content, as the case may be? The question divides into two: what facts determine this for the originating uses of concepts? And what facts are relevant to the content of subsequent uses?

For originating uses of public concepts, we again appeal to theories which answer the analogous question for language. How did the word "Hesperus" come to refer to Hesperus (that is, to Venus)? The intentions of whoever first introduced the word clearly have a role to play. On a simple view, Venus was present to someone's mind, and he coined the word "Hesperus", intending the word to refer to the object in question. The "baptism" was successful, and a practice of using the word for that object was initiated. We can tell the same story for the concept Hesperus, which was presumably originated in the very same event.

A nominative concept may be determined as lacking content in two ways: one is through error, as in the case of the concept Vulcan. The other is through fiction- or myth-making, as was presumably the case for the concept Pegasus (and certainly the case for the concept Sherlock Holmes). In the case of error, nothing satisfied the originating intentions. In the case of fiction, nothing (in our reality) was even supposed to answer to the concept.

There is a great deal more to be said about how the original content of a public concept is fixed but, as we have said, we are happy to borrow from the work of semantic theorists on analogous problems. In the case of concepts without linguistic expression, as in the concepts psychologists attribute to infants and creatures without language, we defer to their expertise. An infant's concept object is determined as having the content it has by a complex interaction among how its uses are caused, with what it covaries, and how it relates to behavior.

These gestures are all we have to say about how original content is fixed. The other question is how it is maintained. Does a concept have its original content for ever, or can it be modified in the course of time? And if it can be modified, how are the modifications effected?

The concept <u>meat</u> used in the seventeenth century had as its reference anything edible; in our terms, its content was the property of *being edible*. The concept <u>meat</u> we use now has as its reference only animal flesh; in our terms its content is the property of *being animal flesh*. One option is to say the earlier concept is the same as the current concept, but that its content has changed. Another option is to say that a new concept, expressed by a word spelled and pronounced the same way, was introduced at some point, and each of the two concepts have retained their original and distinct contents. We prefer the first story, though we think that there are other cases in which the analog of the second story would be preferable.

Originalism entails externalism with respect to the individuation of concepts: thinkers who are intrinsic duplicates may use distinct concepts, for their concepts may originate in numerically distinct events. That's what we are asked to imagine in twin earth cases. A user of the earthly concept <u>water</u> may be an intrinsic duplicate of a user of the twin-earthly concept <u>twater</u>, but they are distinct concepts, originating in different events, on different planets.

Individuative externalism for concepts does not entail semantic externalism for concepts. It might be that though the concept <u>water</u> is distinct from the concept <u>twater</u>, having very different origins, they nonetheless coincide semantically in virtue of their internalistically duplicated use. However, we hold that the right semantics for concepts is externalist, and we take this to be shown by the usual arguments involving twin earth cases. Thinkers who use the concept <u>water</u> may be intrinsic duplicates of thinkers who use the concept <u>twater</u>, but the concepts have different semantics; in our terms, they differ in content, one having water as its content, the other twater. Similarly, the users of the word "arthritis" in Burge's example are intrinsic duplicates, yet they use concepts that differ in their content, one having as its content the property of *being an*

inflammation of the joints, the other the property of *being any inflammation resembling inflammations in joints.*

4. Thoughts

A thought is a well-formed structure of concepts. The thought that Pegasus is a horse is such a structure and so is the thought that Vulcan is a horse. Since one thought contains the concept <u>Vulcan</u> where the other contains the concept <u>Pegasus</u>, they are distinct thoughts. Their distinctness can make all the difference. No one, so far as we know, has (until encountering this paper) entertained the thought that Vulcan is a horse, but the thought that Pegasus is a horse has been widely entertained, thanks to a famous myth.

Atomic concepts combine together in systematic ways. The nonatomic concept <u>the morning star</u> contains the concepts <u>morning</u> and <u>star</u> as proper parts. This nominative concept has a referent, namely Phosphorus, the same referent as the atomic concept <u>Phosphorus</u> and as the atomic concept <u>Hesperus</u>. Whereas the content of the atomic concepts is defined on our account – for both concepts, it is Phosphorus itself – we have not defined the notion of content for nonatomic concepts. Hence, though the concepts <u>the morning star</u> and <u>Hesperus</u> have the same referent, we are not committed to their having the same content.[4]

However, we do define content for a crucial kind of complexes of concepts: thoughts. The structure of a thought can be represented as a tree, in a way familiar from linguistic analysis, but with atomic concepts, rather than words, as the terminal nodes. Originalism requires that there be rules relating the content of whole thoughts to their composing concepts and the contents of these concepts. We are not committed to any specific version of such rules, but we illustrate with ones that associate thoughts with a (possibly empty) set of possible worlds, as a function of the contents (if any) of the constituent atomic concepts, and their mode of combination. On this basis, we extend

---

[4] We could have defined a content for complex nonatomic concepts, but we have found no use for such a notion.

the notion of content from atomic concepts to thoughts: the content of a thought is the associated set of possible worlds.

The association of thoughts with sets of possible worlds presupposes a classification of thoughts by their structures, which will include:

> *unary atomic thoughts:* thoughts consisting just of an atomic nominative concept and a predicatively associated predicative concept, for example the thought that Fido barks.
> *negative thoughts:* thoughts dominated by a concept of negation, for example, the thought that it is not the case that Fido barks.
> *conjunctive thoughts:* thoughts dominated by the concept of conjunction, for example the thought that Fido barks and pigs fly.

The familiar kind of recursion requires the definition of a world-relativized notion of content for atomic concepts. There are various options to choose from, even for just nominative and predicative atoms, but for illustrative purposes we propose:

> *content for atomic nominative and predicative concepts:*
> for any world w, the content of an atomic nominative concept with respect to w is its actual referent, if any; if it has no referent, it has no content at any world.
> for any world w, the content of an atomic predicative concept with respect to w is its actual referent, if any (we will assume this to be the property to which it actually refers, if any); if it has no referent, it has no content at any world.

Here are some examples of the recursive clauses:

- a unary atomic thought is associated with the set W of worlds meeting this condition: $w \in$ W iff both concepts have a content with respect to w, and the content of the nominative concept at w possesses the property that is the content of the predicative concept at w;
- a negative thought is associated with the set W of worlds meeting this condition: $w \in$ W iff w does not belong to the set of worlds associated with the thought that is negated;

- a conjunctive thought is associated with the set W of worlds that is the intersection of the sets associated with each of the conjoined thoughts.

We recognize that thoughts have kinds of complexity going well beyond these simple cases, but dealing with the many other types of combination is not a problem specific to our view: it arises in just the same way for language. In particular, we recognize that there are some constructions, for example thoughts expressing reports of propositional attitudes, in which the content of the whole thought is not determined just by the contents of the atomic constituents. That's what makes it possible for the thought that the Babylonians wondered whether Hesperus is Hesperus to be false, even though the thought that the Babylonians wondered whether Hesperus was Phosphorus is true. The thoughts differ in content, even while having atomic components that agree point-by-point in content.

The *content* of a thought is the associated set of worlds. Every thought has a content. A thought is true iff the actual world is a member of the associated set, and is otherwise false.

As in the case of contents for atomic concepts, the recursive clauses just mentioned are merely illustrative. However, one feature of them is important. According to originalism, concepts are not individuated semantically, so concepts with no semantic content are no embarrassment. This means that we must allow for thoughts to have a content, even if some of their constituent concepts lack contents. This is ensured by the free-logical spirit of the condition for atomic thoughts: if a nominative or predicative concept lacks a referent with respect to the actual world, any atomic thought containing it is false. For example, the unary atomic thought that Vulcan is a planet has a nominative concept with no referent with respect to the actual world, and so no content, and so the set of worlds associated with the thought is empty, so the actual world is not a member of the associated set, so the thought is false. This yields a very straightforward account of true negative existential thoughts, like the thought that Vulcan does not exist. The thought that Vulcan exists is false, for the same reason as the thought that Vulcan is a planet is false; the negation of a false thought is true. So its negation, the thought that Vulcan does not exist, is true. For our theory, the main

requirement is that an atomic concept's failing to have a content should not prevent a thought built from it having a significant role in cognition.

If concepts can change their content, then thoughts can change their truth conditions and so their truth value. Given the content possessed in the seventeenth century by the thought that bread is meat, it was then true; given the content it has now, it is not true. We see nothing problematic in this result. Truth is truth of a thought at a time.

5. Isomorphism

Treating thoughts as structured entities makes it possible to define interesting similarity relations among thoughts, for example:

> *Isomorphism:* Thoughts are *isomorphic* iff they share a complete tree structure, and their corresponding terminal nodes are concepts with the same content.

Isomorphism is sufficient but not necessary for sameness of content. The thought that Odysseus is a Greek is isomorphic to the thought that Odysseus is a Hellene; these thoughts have the same content.[5] The thought that identity is transitive has the same content as the thought that $2 + 2 = 4$, for both are true at every world. They are not isomorphic. The thought that Odysseus is a Hellene is not isomorphic to the thought that the inventor of the Trojan horse is a Hellene, for the latter has a more complex tree structure, preventing any content-preserving one-one correlation of concepts in terminal nodes. (For example, the latter thought will have the concept horse at a terminal node, whereas no terminal node in the former thought will have a concept with the same content.) However, we can define a weaker notion, suggested by the manifest similarity between the pair of thoughts just mentioned. Intuitively, they share an overall subject–predicate structure, and the only difference is that one has an atomic nominative concept where the other has a nonatomic nominative concept with the same referent.

> *Sub-isomorphism:* thoughts are *sub-isomorphic* iff they share a partial tree structure, and each terminal node either corresponds to a coreferential concept

---

[5] This requires that the property of being a Greek is the same property as the property of being a Hellene.

as the corresponding node in the other, or else falls under a higher node that is coreferential with a corresponding terminal node in the other.

The definition ensures that the thought that Odysseus is a Hellene is sub-isomorphic with the thought that the inventor of the Trojan horse is a Hellene.

We can also define a more demanding notion, in which thoughts related by it also match with respect to their internal patterns:

Super-isomorphism: thoughts are *super-isomorphic* iff they are isomorphic and if any two terminal nodes in one contain the same concept so do the corresponding terminal nodes in the other.

The thought that Hesperus is Hesperus is isomorphic to, but not super-isomorphic to, the thought that Hesperus is Phosphorus. It is super-isomorphic to the thought that Phosphorus is Phosphorus.

In applying originalism to puzzling cases, and especially in connection with the adequacy of reports of beliefs and other thoughts, such structural relations play a significant role.

6. Indexicality

Is there a single concept I, used both by you and by us? If we say there is a single concept, we are forced to conclude that one concept has many different referents, even at a single time; we would be landed with ambiguity on a massive scale. Saying that there is a single concept is also inconsistent with originalism: different subjects introduce their I-concepts on different occasions. We introduced ours in the twentieth century; Hume introduced his in the eighteenth. So the concepts are distinct.

Accordingly we prefer the alternative view: there is, strictly speaking, no single concept I we all share. We'll speak of the *concept-template* I, contrasting this with a given subject's specific concept I. Only specific concepts are concepts. A concept-template is not a concept, but rather a recipe for forming concepts, or a pigeon-hole to hold many different but similar concepts.

The thought you would express by the words "I am hungry" is distinct from the thought that someone else would express with those words; the thoughts are not even

isomorphs, since their first elements differ in content. The thought that you are hungry gets closer, for it is an isomorph of the thought you think. But we take it that your concept I was introduced on a different (and much earlier) occasion from that on which we introduced a specific concept you for you, so the concepts are distinct. Hence the thought that we think when we think that you are hungry is not the same thought as the one you thought. We have reached a somewhat Fregean position from a quite different starting point.[6]

Likewise for other indexicals: there is a demonstrative concept-template that, and many specific concepts that, used on specific occasions with specific intentions. Questions of content arise only for specific concepts; the notion does not apply to concept-templates.

It is one thing to reuse a specific demonstrative concept that, another to introduce a new specific demonstrative concept that. A subject may use specific that concepts quite independently, for manifestly different objects. In this case, we have more than one specific concept forged from a single concept-template. But a subject may also exercise the very same specific that concept more than once. These latter, non-originating cases, are marked by deference to previous uses; this feature is absent from the former, originating uses.

Gareth Evans uses a case involving indexicals (owed to John Perry) to support Fregean views:

> Suppose a person can see two views of what is in fact one very long ship, through two windows in the room in which he is sitting. He may be prepared to accept "That ship was built in Japan" (pointing through one window), but not prepared to accept "That ship was built in Japan" (pointing through the other window). Now suppose we try to describe this situation in terms of the ordered-couple conception of Russellian thought. We have a single proposition or thought

---

[6] Frege says that "everyone is presented to himself in a special and primitive way, in which he is presented to no-one else" (1918: 359). In comparing our view with Frege's, one should not rush to equate our thoughts with Fregean Gedanken. The elements of a Fregean Gedanke are individuated semantically.

content – <the ship in question, the property of having been built in Japan> – to which the subject both has and fails to have the relation corresponding to the notion of belief. Not only does this fail to give any intelligible characterization of the subject's state of mind; it appears to be actually contradictory (Evans 1982: 84).

The example helps bring out the difference between traditional views and our own. The two uses of the complex demonstrative concept-template that ship involve distinct specific demonstrative concepts, one introduced in connection with the first sighting of the ship, the other introduced in connection with the second. This can be inferred from the speaker's intentions and reactions. For example, he has no inclination to bring forward the information *was built in Japan* when having the thought associated with the view from the second window. Hence we approach a description of the case with two concepts to attribute to the thinker, concepts originating in distinct acts of concept-introduction.[7]

Once we have distinct concepts, no further difference is needed to "give an intelligible characterization of the subject's state of mind", for distinctness of concepts yields distinctness of thoughts. We ascribe the same content to the thought the subject accepts as to the thought he rejects, just as does the "ordered couple" conception that Evans attacks. Both specific demonstrative concepts have the same content. Hence the set of worlds at which the first thought is true (the thought expressed by the first utterance of the words "That ship was built in Japan", and which the thinker accepts) is the same as the set of worlds at which the second thought is true (the thought the thinker rejects). Subjects do not relate directly to sets of worlds. We can't infer that the set of worlds in question has the contradictory properties of being both accepted and rejected by the subject; sets of worlds are not the kinds of thing that can be accepted or rejected. Rather, we relate to sets of worlds via thoughts, and there are distinct thoughts. The situation is exactly like one in which a subject is prepared to accept the

---

[7] It's a feature of indexical concepts that a speaker can introduce them, in a quite simple way. This contrasts with public concepts acquired by immersion, like the concept Paderewski. It's not up to us to determine anything about the nature or semantics of that concept.

thought expressed by "Hesperus is visible" but is not prepared to accept the thought expressed by "Phosphorus is visible". The explanation is based on conceptual difference, not semantic difference.[8]

Cases like the ship seen through different windows are problematic for Fregeans, since applying the standard Fregean strategy requires uncovering distinct senses, one but not the other of which the subject is prepared to accept. But it's hard to explain what the different senses of "that ship" might be. By contrast, originalists have no such difficulty. We do indeed need two of something, but we have two specific demonstrative concepts, and their distinctness is characterized in the originalist way: they are concepts with different origins.

7. Cognition

A Fregean datum is that it's one thing to think that Hesperus is Hesperus, and another to think that Hesperus is Phosphorus; one thing to think that Hesperus is visible, another to think that Phosphorus is visible. We agree. Different thoughts are involved, that is, different structures of concepts, since the concept <u>Hesperus</u> is distinct from the concept <u>Phosphorus</u>. We disagree with Fregeans that the difference requires postulating any additional semantic layer. As illustrated in the example from Evans of the ship built in Japan, we think the differences can be fully and satisfyingly explained using just concepts (and their combination into thoughts) and their contents (even when these are conceived as semantically very coarse-grained).

Distinct concepts can, and typically will, play different roles in our cognitive activities, even if they have the same content. This can be so for more than one reason. One is that a thinker may not know that concepts she uses have the same content, even if they do, and even if she knows the content of each. This is how we can explain an ancient (pre-Babylonian) astronomer believing that Hesperus is visible but not that Phosphorus is visible, and our ship-viewer as believing that that ship, but not that ship, was built in Japan. Another role for concepts with the same content is in cases in which, although a

---

[8] Ruth Millikan (e.g. 2000: ch. 12) has also stressed the importance to cognition of sameness and difference among vehicles of content, as opposed sameness and difference of content.

thinker knows quite well, concerning two concepts, that they have the same content, she may attain a more accurate report of the beliefs of others by using one concept rather than another. That's what we just did in reminding you about the ancient astronomers. A third role for concepts in cognition is that their sameness and difference generate relational effects: the pattern exemplified by the thought that Hesperus is Hesperus is distinct from the pattern exemplified by the thought that Hesperus is Phosphorus. In the first, a single concept is used twice. In the second, two concepts are each used once. This may have an impact on informativeness. When a concept is used twice in an identity thought, the thought is typically uninformative, whereas a thought similar in point of structure and content, but using two concepts, may be informative.

To vary the example: the thought that Greeks are Greeks is typically uninformative, whereas the thought that Greeks are Hellenes is potentially informative. In processing the first, only one concept is exercised, though on two occasions. Whatever processing effort is required has already been made by the time the second occurrence of the concept is encountered; the previous interpretive outcome can simply be brought forward. In processing the second thought, more effort is required, since a second concept needs to be processed from scratch. Corresponding to the extra effort, there is more informational value. To use a computer analogy: a comparison is made between distinct addresses, and the need for such a comparison is independent of the values held at the addresses.

This difference in degree of informativeness can be made to affect truth value, if a suitable embedding is chosen. For example, the first of the following pair is true, the second false:

1. The thought that Greeks are Greeks is exactly as informative as the thought that Greeks are Greeks.
2. The thought that Greeks are Greeks is exactly as informative as the thought that Greeks are Hellenes.

This analysis shows that what we have lumped together as Fregean data come in at least two kinds. There's difference of informativeness that depends on difference in the patterning of concepts within a single thought, and not on the intrinsic features of the

concepts themselves. Then there are differences between thoughts that differ only in their concepts, and not in their patterning or their content, like the difference between believing that Hesperus is visible and believing that Phosphorus is visible. Super-isomorphism preserves the internal patterning, so super-isomorphic thoughts should be equally informative. Isomorphic thoughts may not be. The originalist explanation has a degree of detail and accuracy that cannot be matched by other theories.

We claim that the work supposedly done by difference of sense can be done better by difference of concept. Unlike sense, which raises tricky questions about its identity and its metaphysical status, concepts have clear identity conditions based on their origin and are, we think, required by any reasonable account of the mind. Hence we think one should not appeal to mysterious difference of sense when one can instead appeal to the clearer notion of difference of concept.

One problem that we regard as especially difficult for Fregeans is posed by Mates cases. One can adapt essentially Mates's argument to show that no two expressions have the same sense. That must strike the Fregean as an extraordinary conclusion. What could stop two expressions agreeing in sense? Originalism can explain the puzzling character of Mates cases. Pablo can wonder whether everyone who believes that Greeks are Greeks believes that Greeks are Hellenes because the concept Greeks is distinct from the concept Hellenes.[9] Hence Pablo is wondering whether everyone who believes one of two distinct thoughts believes the other. This is evidently unproblematic.

The upshot of this section is that the nature of cognition depends, unsurprisingly, not only on contents but also on the vehicles which serve to represent contents: concepts and thoughts. These are metaphysically real elements of our mental life, involved in reasoning and related cognitive activities. Concepts and thoughts can differ, and thereby differentially affect cognition and action, even when they have the same content.

---

[9] The Illyrians introduced the concept Greek (using a word resembling "Greek") for Greeks when they first encountered them (in Italy), but at that point the Greeks already thought of themselves as Hellenes (using a much older concept).

8. "Mastering" or "grasping" concepts

Traditional theories often make no distinction between understanding a concept, grasping it, being able to use it correctly, mastering it, and knowing its content. For us, by contrast, there are important distinctions. There's what it takes to acquire an atomic concept. There's what it takes to combine atomic concepts into complex structures, for example thoughts. There's what it is to know the content of a concept, if it has one, or to know that it has no content, and this takes different forms in atomic and nonatomic cases. And there's what it is to know the truth conditions of thoughts. Since some concepts lack contents, yet can be used correctly, we must firmly distinguish between being able to use a concept correctly and knowing its content (if this last phrase is so understood that the knowledge requires the concept to have a content).

There are three main ways in which concepts can be acquired: (i) by subpersonal processes, as part of normal development; (ii) by explicit acts of concept introduction; and (iii) by immersion in a concept-using practice. We acknowledge that there is work to be done in saying more about these modes of acquisition. But we think their reality cannot be doubted. We think it highly unlikely that any analysis of what these ways involve would in general require mention of knowledge of their content (in our sense of "content").

To defend our position, consider a rival view committed to principles like these:

> correct use of a nominative concept requires a disposition to apply it, under
> favorable circumstances, to the object to which it refers

A rival theorist might suggest that a subject with such dispositions thereby knows the content of the concepts involved; hence knowledge of content is implicated in a proper account of concept possession.

It's easy to see that this kind of principle can't be right as stated. First, consider nonatomic nominative concepts, like the concept the present King of France. There is no object to which it refers, so someone who gasps the concept cannot in fact apply it, or be disposed to apply it, to the object to which it refers. Even were there such an

object, a thinker might never have given it a moment's thought, or never conceptualized it in any such way.

A first modification restricts such principles to atomic concepts. The empty case remains problematic. It's unclear that Le Verrier was disposed to apply the concept <u>Vulcan</u> to anything; but if he had had such a disposition it would have been a manifestation of error, and so not a basis for knowledge of content (or of anything else). The concept <u>Vulcan</u> has no content, so mastering it cannot require knowledge of content. On some views about properties, the same can be said for the concept <u>phlogiston</u>.

Such principles are problematic independently of empty cases. The young Burge, we know on his own authority, was not disposed to apply the concept <u>fortnight</u> to the things possessing the property it expresses, namely periods of fourteen days. Fodor tells us that he was at one point unable tell an elm from an oak, and so his dispositions to apply the concept <u>oak</u> do not differentiate his use of this concept from his dispositions to apply the concept <u>elm</u>.

More generally, concept possession is consistent with all sorts of mistakes and misunderstandings about the concept's subject matter (whales are fish, there are many witches, and so on). Perhaps not just any false belief is consistent with possession of an atomic concept (it is difficult to say anything general about when error prevents concept possession and when it does not), but the possibility of error is widespread enough to undermine attempts to forge a general link between concept possession and knowledge.

The general claim is consistent with there being exceptions, special cases in which understanding a concept requires knowledge of its content. Here are two candidates for being such exceptions: specific indexical concepts; and logical concepts. Suppose a thinker introduces a new specific indexical concept <u>that</u>. Is not its referent, and so its content, determined by the thinker's intentions? And must not the thinker know what these intentions are, and so know the concept's content? The suggestion is confused. Even supposing the referent is determined by intentions, there is no guarantee that there will be a referent. There may even be a guarantee that there is no referent, as

when a parent in a game asks a child to guess what *that* is, holding out a closed fist the parent knows to contain nothing. (The parent may continue: is that a sweetie?, and so on, making plain that a usable concept has been introduced.) And even if there is a referent, knowledge of the intention will not guarantee knowledge of the referent, in every contextually salient interpretation of this demand. In a phonetic identity parade, we are asked to identify our attacker by his voice (the attack took place in the dark, and the police are anxious to ensure we are not prejudiced by visual features of the suspects). The two candidates are asked to say something, while hidden from us by a screen. We recognize the voice of the speaker on our left, and proclaim "That's the man!". Before the curtain is lifted, the men switch places. We believe the referent of the demonstrative concept <u>that</u> which we exercised in the corresponding thought was the man now on our left, whereas it was the man now on our right. So, while knowing our intentions, we mistake the content of a concept we introduced.[10]

Conjunction is a prime example of a logical concept of which it seems plausible to say that grasping it ensures or requires knowledge. Perhaps I understand this concept iff[11] I know that if a conjunction is true, so are the conjuncts, and if two thoughts are true, so is their conjunction. But suppose a subject who appears to possess a concept of conjunction thinks this is not quite right: order can make a difference, whereas the conditions just stated preclude this.[12] To us, that appears to be a mistake *about conjunction*, that is, a mistake on the part of one who genuinely uses the concept of conjunction. On the alternative view, according to which the subject has some related but distinct concept, there is no mistake at all, and so nothing a sensible person could disagree with.

---

[10] A better way of appealing to indexicality as imposing a knowledge requirement on concept use highlights the analog, for concepts, of character (in Kaplan's sense). We address this in forthcoming work.

[11] Given that what follows holds (on many views) for the distinct concept <u>but</u>, "only if" might be better than "iff".

[12] We ourselves are inclined towards this view (to which update semantics are sensitive). But the example treats it as a mistaken view.

Even though knowledge of content is typically not required in order to use an atomic concept, thinkers often do know the contents of their concepts. To know the content of a given concept is to know *what* the content of the concept is, and this in turn is being able to give a correct answer to the indirect question introduced by "what". However, we should not require that typical subjects possess the concept content, for that is a technical notion that thinkers who have relevant knowledge may lack. Hence we cannot rely on thinkers to have a good answer to a question conceptualized using the concept content, even though we intuitively wish to credit them with some suitably related knowledge.

One way to meet this difficulty is to suggest that if a thinker knows that the concept Hesperus refers to Hesperus, she knows what the content of the concept Hesperus is. Here we use, and require the knowing subject to use, the non-technical notion of reference, a notion on the basis of which the more general and technical notion of content was introduced. Even if this is an indisputable sufficient condition, it should not be considered necessary: thinkers can use concepts without having the concept reference (young children are examples of such thinkers).

Another way to meet the difficulty is to say that a thinker counts as knowing the content of a concept if she is disposed to use it of things to which it applies. Because of empty concepts (among other reasons), this could at best be a sufficient, and not a necessary condition. It has the advantage of not requiring those who know the content of their concepts to have any semantic concepts.

The two approaches correspond to two ways of thinking of knowledge of content. On the first approach, it is something theoretical, manifest in knowledge of facts linking concepts and the world, and so available only to possessors of semantic concepts. On the second approach, it is more behavioral, inferred from a kind of practical ability. The approaches are not in conflict: one could develop each to provide accounts of both theoretical and practical knowledge of content. As noted, neither account provides a necessary condition for what it is to possess a concept. From the originalist perspective, the analysis of knowledge of conceptual content promises no immediate rewards in terms of applications to puzzling cases.

Understanding and knowledge relate differently to concepts and to thoughts. In a long tradition, understanding sentences has been associated with knowing their truth conditions. On our view, a thought is true iff the actual world belongs to its associated the set of worlds. Hence we have the resources to identify knowing the content of a thought with knowing such a truth condition, in conformity with the traditional view. This makes for a contrast between concepts and thoughts: thinking involves knowing truth conditions, but possessing a concept typically does not require comparable knowledge.

The justification for the traditional requirement derives from what is involved in thinking. A thinker must possess all the concepts in the thought's terminal nodes. The requirement that is harder to articulate is that the thinker assemble these elements into an appropriate tree structure. This is where we see the need for something like knowledge of the semantic axioms that govern concepts, leading to knowledge of truth conditions. The concepts composing a thought should not simply be present together: they need to be structured in accordance with suitable rules, rules that will determine how the truth conditions of the whole thought emerge from the constituent concepts and their contents. The simplest way to model how this is done is in terms of the subject's applying the right semantic rules to the atomic concepts.

The simple model is unsatisfactory in at least two ways. First, it attributes to thinkers much more sophistication than they need have in order to think. It's obvious that young children do not have semantic concepts, and could not grasp semantic rules. Indeed, very few fully realistic semantic rules for natural languages have been explicitly formulated, so far as we are aware. Second, there is something unsatisfactorily regressive about the model: applying a rule is thinking, but if every act of thinking needs a prior act of thinking, the process can never get started. We therefore have to appeal to a model of some less demanding kind.

The topic has been widely discussed in the context of language and, once again, we propose to take over whatever theory is best.[13] As a familiar label, we think that what is required is "tacit knowledge" of the semantic axioms: a thinker must incorporate a

---

[13] See Davies 1987, Wright 1986.

processing system that operates, in a subpersonal way, rather as a personal system would operate in applying explicitly known semantic rules. The subpersonal operations are not acts at all, and so are not acts of thinking; the regress is removed.

9. *Conclusion*

Our theory combines the best features of Fregean and Millian views. Fregeans should respond favorably to our making room for the data they say should be taken at face value, as well as making room for intuitively similar data that they are unable to accommodate (as in demonstrative and Mates cases). Millians should respond favorably to our one-level view of content. From our perspective, Fregeans are right to think that something more than reference is needed in a complete account, but wrong to think that this something more needs to be epistemically or semantically individuated. Millians are right to think that content is referential, but wrong to think that nothing else is needed to explain cognition. Cognitive processing depends on the vehicles of content, concepts and thoughts, not just on their content.[14]

For the application of originalism to a range of puzzles, together with further development of the theory, see our forthcoming *Seven Puzzles of Thought and How to Solve Them: An Originalist Theory of Concepts*, Oxford University Press.

REFERENCES

Burge, T. (1979). "Individualism and the mental". *Midwest Studies in Philosophy IV: Metaphysics.* P. A. French, T. E. Uehling and H. K. Wettstein (eds). Minneapolis, University of Minnesota Press: 73–121.

Davies, M. (1987). "Tacit knowledge and semantic theory: Can a 5% difference matter?" *Mind* 1987: 441–62.

Evans, G. (1982). *The Varieties of Reference*. Oxford, Clarendon Press.

Frege, G. (1918). "Logical investigations: thoughts." *Collected Papers on Mathematics, Logic and Philosophy*. B. McGuinness (ed.), Oxford, Basil Blackwell. 1984: 351-72.

Kripke, S. (1980). *Naming and Necessity*. Oxford, Basil Blackwell.

Mates, B. (1950). "Synonymity." *University of California Publications in Philosophy* 25: 201–26.

Millikan, R. (2000). *On Clear and Confused Ideas*. Cambridge University Press.

Pylyshyn, Z. (2007). *Things and Places: How the Mind Connects with the World*, MIT Press.

Tye, Michael (2005). "Nonconceptual content, richness, and fineness of grain." In T. Gendler & J. Hawthorne (eds.) *Perceptual Experience*, Oxford University Press.

Tye, Michael (forthcoming). "Attention, seeing and change blindness." *Philosophical Issues*.

Wright, C. (1986). "Theories of meaning and speakers' knowledge". *Philosophy in Britain Today*. S. G. Shanker (ed. London and Sydney, Croom Helm: 267–307.