Analyticity: the Carnap-Quine debate and its aftermath

Gary Ebbs

Draft 4/25/2018
Forthcoming in Becker and Thomson, eds.,
Cambridge History of Philosophy: 1945-2015

W.V. Quine's criticisms of Rudolf Carnap's efforts to draw a boundary between analytic and synthetic sentences shook mid-20th century Anglo-American philosophy to its foundations, leaving logical empiricism in ruins, and sparking the development of radically new ways of theorizing that continue to shape Anglo-American philosophy today. Despite decades of discussion, however, neither Carnap's analytic-synthetic distinction nor Quine's criticisms of it are well understood. My central goals here are to summarize and clarify them, evaluate influential objections to Quine's criticisms, survey related work on analyticity and apriority by Hilary Putnam, Saul Kripke, David Chalmers, Paul Boghossian, and Gillian Russell, among others, and briefly discuss whether meaning is determinate in ways that recent explications of analyticity require.

1. How analyticity came to be regarded as central to philosophy

Since at least as early as Kant, philosophers have distinguished between analytic statements, which they suppose to be true solely by virtue of what the statements mean, and synthetic statements, which they suppose to be true, if true at all, by virtue of both what the statements mean and of extralinguistic facts. Kant's version of this distinction focuses not on statements, but judgments, and, in particular, on judgments "in which the relation of a subject to the predicate is thought" (Kant 1781, p. 48). A judgment of this form is analytic, according to Kant, if the predicate is "contained" in the subject; otherwise the judgment is synthetic. Kant argued that one can know a priori, i.e., independent of empirical evidence, that an analytic judgment is true. He also argued that our knowledge of arithmetic and geometry is a priori. Limited by the traditional syllogistic logic of his day, however, he saw no way to view such knowledge as analytic. Even the simplest axioms of arithmetic, such as "Every number has a successor," resist paraphrase in the monadic quantificational terms of traditional syllogistic logic (Friedman 2010, p. 591). Kant concluded that our knowledge of arithmetic and geometry is synthetic a priori. The leading question in Kant's *Critique of Pure Reason* (Kant 1781) is "How is synthetic a priori knowledge possible?" His revolutionary answer implies that much of traditional philosophy is misguided and fruitless, the result of efforts to answer questions that transcend reason's powers. Kant labeled all such efforts "metaphysics."

Kant's view that arithmetic and geometry are synthetic a priori was challenged and gradually discredited by a series of advances in logic, geometry, physics, and philosophy in the 19th and early 20th centuries (Friedman 2010). Building on these advances, in the 1910s and 1920s a group of scientifically-oriented philosophers, including Rudolf Carnap, Hans Reichenbach, and Moritz Schlick, devised ways of formulating Albert Einstein's revolutionary theories of special and general relativity in terms of a new conception of analyticity that encompasses polyadic quantificational logic and axioms of arithmetic and even set theory, as needed, and an exclusively empiricist

conception of synthetic statements. Contrary to Kant, these new philosophers, the logical empiricists (also known as the logical positivists), concluded that there are no synthetic a priori truths. They nevertheless endorsed the spirit of Kant's critique of traditional metaphysics: just as Kant dismissed any question that transcends (what he regarded as) reason's powers as misguided and fruitless, so the logical empiricists dismissed any supposed question that cannot be formulated in a well-defined formal system with empirical applications as devoid of cognitive content.

2. Carnap's analytic-synthetic distinction

The logical empiricist's efforts to clarify their conception of cognitively meaningful discourse culminated in Carnap's work in the 1930s and 1940s. Starting in *Logical Syntax of Language* (Carnap 1937, henceforth *Syntax*), Carnap fashioned a new form of logical empiricism, *Wissenschaftslogik*, or the logic of science, rooted in the principles that everything that can be said is said by science (Carnap 1934, p. 47) and that if investigators are to agree or disagree at all, they must share clear, explicit criteria for evaluating their assertions (Carnap 1963a, p. 44–45). Carnap rejects both rationalism and empiricism about logical and mathematical truths. To explain how a logical or a mathematical truth can be acceptable yet not justified by a rational perception or empirical evidence, Carnap formulates such truths so that they are analytic, or "true in virtue of meaning" (Carnap 1963a, pp. 46–47, 63–64).

Carnap's development of these ideas in *Syntax* is inspired and shaped by Gödel's incompleteness theorems (Gödel 1931). Using methods Gödel devised to prove his new theorems, Carnap shows how to construct syntactical definitions of 'logical consequence' for what he called language systems—artificial languages with built-in syntactical and semantical rules.[1] Each language system *LS* is characterized by its *formation rules*, which specify the sentences of *LS*, and its *transformation rules*, which together settle, for every sentence *s* of *LS* and every set *R* of sentences of *LS*, whether or not *s* is a *logical consequence* of *R* in *LS*.

A typical language system contains *finite* transformation rules—primitive sentences and rules of inference each of which refers to a finite number of premises—otherwise known as rules of deduction. Gödel's first incompleteness theorem shows that if *LS* contains elementary arithmetic, the rules of deduction of *LS* do not together define the logical consequence for *LS*. Carnap therefore proposed that we define the logical consequence relation for such languages either by laying down *transfinite* rules of transformation (also known as omega rules), which are defined for an infinite number of premises, or by adopting a syntactical definition of truth for the logical and mathematical

[1]Soon after publishing *Syntax* Carnap learned of Tarski's method for defining truth for formalized languages (Tarski 1936), and he began including Tarski-style truth definitions among the rules of a language system. He noted that a syntactical definition of truth for the logical and mathematical parts of a language system "can be brought into a technically more simple but not essentially different form by a procedure … (originated by Tarski) of defining 'true' in semantics" (Carnap 1942, p. 247). In the exposition that follows this note, the phrase "transformation rule" should be understood to encompass both syntactical and Tarski-style semantical rules of a given constructed language system.

parts of a language system (Carnap 1937, §§14, 34, 43–45). It is in terms of the logical consequence relation of a given language system that Carnap defines "analytic", "contradictory", and "synthetic" as follows:

> A sentence *s* is *L-valid* (*analytic)* in *LS* if and only if the L-rules of *LS* together settle that *s* is a logical consequence in *LS* of the empty set of sentences of *LS*;
>
> *s* is *L-contravalid* (*contradictory)* in *LS* if and only if the L-rules of *LS* together settle that the negation of *s* is a logical consequence of the empty set of sentences of *LS* (Carnap 1937, §52).
>
> *s* is *synthetic* in *LS* if and only if *s* is neither analytic nor contradictory in *LS.*

A language system suitable for empirical science also contains synthetic sentences, some of which, the so-called *protocol* sentences, such as "That spot is red", "This piece of paper is blue", are classified by their syntactical form as suitable for expressing empirical observations. (Carnap 1936–1937, pp. 454–456; and Carnap 1937, p. 317)

In *pure* syntax or semantics we may specify any language system we please and investigate its logical consequences in abstraction from any actual language. In *descriptive* syntax and semantics, we may stipulate that the sentence forms of a language system we have constructed in pure syntax or semantics are to be correlated with particular strings of sounds or marks that we can use to make claims (Carnap 1942, §5). Once we have stipulated such correlations, a sentence of our physics, for instance,

> will be *tested* by deducing consequences on the basis of the transformation rules of the language, until finally sentences of the form of protocol-sentences are reached. These will then be compared with the protocol-sentences which have actually been stated and either confirmed or refuted by them. (Carnap 1937, p. 317)

Carnap emphasized that although our decision as to which sentences of physics state laws of nature is based upon the observations made so far, "nevertheless it is not uniquely determined by them" (Carnap 1936–1937, p. 426). He concluded that such decisions are partly "conventional" despite their "subordination to empirical control by means of the protocol sentences" (Carnap 1937, p. 320). In short, "the [physical] laws [of a language system] are not inferred from protocol sentences, but are selected and laid down" in accord with one's commitments regarding the existing protocol sentences (Carnap 1937, pp. 317–18).

To make explicit the conventional element in our decisions as to which physical laws to accept, and thereby forestall unscientific questions about whether our acceptance of such laws is justified, Carnap proposes that we build physical laws into a language system by stipulating P-rules, or physical rules, for the system (Carnap 1937, §51). With P-rules in place, we can define P-validity as follows:

> Sentence s is *P-valid* in language system *LS* if and only if *s* is a consequence in *LS* of the set of L- and P-rules of *LS* and *s* is not L-valid (analytic) in *LS*.

The key methodological significance for Carnap of this and the previous definitions is that if a sentence of a language system is L-valid or P-valid then *(first) anyone who has chosen to use the system, understands its rules, and has derived the sentence from the rules is thereby committed to accepting the sentence; and (second) there is no legitimate 'higher' or 'firmer' criterion for judging whether the sentence is true* (Carnap 1934, p. 46).

The rules for language systems are specified in such a way that they can be regarded as part of science—the science of arithmetic—via Gödel numbering and recursion theory. Since mathematicians freely propose and derive consequences from any definitions they like, and logical syntax can be viewed as part of arithmetic, it is in accord with scientific practice to adopt any formation and transformation rules we like (Carnap 1937, §17). This attitude is encapsulated in Carnap's Principle of Tolerance:

> *In logic, there are no morals*. Everyone is at liberty to build up his own logic, i.e. his own form of language, as he wishes. All that is required of him is that, if he wishes to discuss it, he must state his methods clearly…. (Carnap 1937, p. 52).

When combined with Pierre Duhem's observation that a statement has empirical consequences only if it is conjoined with other statements, Carnap's Principle of Tolerance engenders a radically permissive view of how rules of a language system may be revised in the face of unexpected observations:

> If a sentence which is an L-consequence of certain P-primitive sentences contradicts a sentence which has been stated as a protocol-sentence, then some change must be made in the system. For instance, the P-rules can be altered in such a way that those particular primitive sentences are no longer valid; or the protocol-sentence can be taken as being non-valid; or again the *L*-rules which have been used in the deduction can also be changed. *There are no established rules for the kind of change which must be made*. … All rules are laid down with the reservation that they may be altered as soon as it is expedient to do so. This applies not only to the P-rules but also to the L-rules, including those of mathematics. (Carnap 1937, pp. 317–318, my emphasis)

Although we are committed more firmly to some rules of a system than we are to other rules of the system, according to Carnap, "In this regard there are only differences of degree; certain rules are more difficult to renounce than others" (Carnap 1937, p. 318).

3. Quine's criticisms of Carnap's analytic-synthetic distinction

Quine travelled to Prague in the early 1930s to meet Carnap, thus beginning a decades-long intellectual exchange with him (Quine 1986, pp. 11–13). In "Truth by Convention" (Quine 1936) Quine explains that logic cannot be true by explicit convention, since there are infinitely many logical truths and one needs to presuppose logic to derive those truths from any set of finitely many explicit conventions. This observation supplements Carnap's earlier observation in *Syntax* that due to a syntactical version of Tarski's Undefinability Theorem, if *LS* is a consistent language system rich enough to express elementary arithmetic, then "analytic-in-*LS*" cannot be defined in *LS*, and therefore depends for its formulation and application on the logic of a stronger metalanguage (Carnap 1937, Theorem 60C.1; Ebbs 2011a). These technical points, on which Carnap was clear, make plain that Carnap does not hold that the logical or mathematical truths of a given language system are "made true" by its transformation rules (as Boghossian 1996, p. 365, claims), or that such truths are a priori in a traditional sense (as Sober 2000, p. 259–260, and Soames 2003, p. 264, claim).

Quine's goal in "Two Dogmas of Empiricism" (Quine 1953a, henceforth "Two Dogmas") is to demonstrate that "empiricists," especially Carnap, have failed to explain their term 'analytic' in a language suited for and used in the sciences. As Quine later explained, his doubts about Carnap's analytic-synthetic distinction were an expression of "the same sort of attitude, the sort of discipline that Carnap shared and that I owed, certainly, in part to Carnap's influence: I was just being more carnapian than Carnap in being criticial in this question" (Quine 1994, p. 228).

Quine places two constraints on a successful explication of 'analytic': (first) sentences that are analytic "by general philosophical acclaim" ("Two Dogmas," p. 22) should be analytic according to the explication; and (second) a sentence should be analytic according to the explication only if it is true ("Two Dogmas," p. 34). He assumes that first-order logical truths, defined as true sentences that "remain true under all reinterpretations of [their] components other than the logical particles," are among the acclaimed "analytic" sentences ("Two Dogmas," pp. 22–23). In taking this first step, Quine follows Carnap, who proposed it himself in discussions with Tarski and Quine in 1941 (Frost-Arnold 2013, p. 156). Quine then proposes that a sentence is in the wider class of analytic truths if "it can be turned into a logical truth by putting synonyms for synonyms" ("Two Dogmas," p. 23), and notes that to specify the wider class of analytic truths, we need a precise, unambiguous notion of synonymy. In taking this second step, Quine again follows Carnap, who writes, "A precise account of the meaning of the L-terms [including "L-true" (analytic)] has to be given by definitions for them" (Carnap 1942, p. 62; see also Carnap 1963c, p. 918).

Quine's question in §2 of "Two Dogmas" is whether the wider class of analytic truths can be regarded as true by definition. His answer is that with the sole exception of explicitly adopted definitional abbreviations, the synonymies we would need to clarify in order to specify the wider class of analytic truths are not "created" by definitions, but presuppose independent and prior relations of synonymy that do not hinge on our adoption of definitions (Ebbs 2017b).

In §3 of "Two Dogmas," Quine considers whether one can define synonymy for two linguistic expressions in terms of their inter-substitutivity in all logically relevant grammatical contexts without a change in truth-value. He points out that this strategy will satisfy the above constraints on a successful explication of 'analytic' only if inter-substitutivity is defined relative to a language that contains the word 'necessarily'. The strategy is therefore ultimately circular, if, as he, Carnap, and other scientific philosophers assumed, modal notions such as necessity and possibility are themselves unclear, and should be explicated, if at all, in terms of analyticity.

Carnap can agree with all of the arguments in §§1-3 of "Two Dogmas". The main point of these sections is to clear the ground for Quine's criticism, in §4, of Carnap's method of drawing an analytic-synthetic distinction. Quine's criticism is that Carnap's methods of defining analyticity allow us to classify the sentences of a given natural language as "analytic" or not only relative to a conventional decision about which language system corresponds with, or describes, the natural language. (Recall that it is the task of *descriptive* syntax and semantics to specify such correlations.) Even assuming that analytic sentences are supposed to be true, by deciding to label sentences "analytic-in-L_1", "analytic-in-L_2", etc., we do not explain "analytic in L" for variable L. In short, Quine argues, Carnap's methods fail to explain what the labels "analytic-in-L_1", "analytic-in-L_2", mean.

This difficulty is compounded by Carnap's distinction between L-rules and P-rules. As noted in §2 above, Carnap stresses that

> all rules are laid down with the reservation that they may be altered as soon as it is expedient to do so. This applies not only to the P-rules but also to the L-rules, including those of mathematics (Carnap 1937, p. 318).

From this methodological point of view, the distinction between L-valid (analytic) or P-valid sentences, hence also the distinction between analytic and synthetic sentences, looks arbitrary (Quine 1963, p. 398; Hempel 1963, §VII; Friedman 2010, pp. 671–673). To make matters worse, when Quine wrote "Two Dogmas," Carnap had still not explained how to draw the distinction to his *own* satisfaction[2], and none of the explications of it that Carnap later proposed, many of which (e.g., Carnap 1956; Carnap 1963a, p. 964) are complicated and technical, provide a language-system-independent criterion for classifying a sentence as L-valid (analytic), not P-valid. (For a carnapian reply to these concerns, see Friedman 2010, pp. 673–678.)

In §5 of "Two Dogmas," Quine argues that one cannot draw a boundary between analytic and synthetic sentences directly in verificationist terms. Following Duhem and

[2]In 1966, Carnap writes, "although I did not share the pessimism of Quine and Hempel [about the prospects of drawing an analytic-synthetic distinction for theoretical vocabulary] I always admitted that it was a serious problem and that I could not see a satisfactory solution. . . . Finally, after many years of searching, I found this new approach [in Carnap 1956 and Carnap 1963a, p. 964]." (Carnap 1966, pp. 276–277)

Carnap (1937, p. 318, cited above), Quine argues that since no sentence can be tested by experience in isolation from other sentences, one cannot define analyticity directly in terms of confirmation by experiences. Some readers (e.g. Hookway, pp. 36–37) have taken this argument as a criticism of Carnap. In Carnap's view, however, the analytic-synthetic distinction, including the relationships between theoretical sentences and protocol sentences, can be drawn only relative to a language system defined by explicit formation and transformation rules. Since Quine assumes in §5 that there are no such rules, his reasoning in §5 does not directly engage or conflict with Carnap's views.

In §6 of "Two Dogmas" Quine recommends that we view science as a "field of force whose boundary conditions are experience" ("Two Dogmas," p. 42). He sees no need to draw a boundary between analytic and synthetic sentences of a language. It is enough to apply scientific method, and to note that some statements are more firmly accepted and accordingly less likely to be revised in the face of surprising new experiences than others. In this way Quine aims to purge Carnap's terms "L-rule", "P-rule", "analytic" and "synthetic", from his description of scientific methodology, while remaining true to the radical spirit of Carnap's scientific philosophy, by rejecting the traditional philosopher's goal of providing justifications that are 'higher' or 'firmer' than any that science provides.

4. Grice and Strawson in Defense of Analyticity

In "In Defense of a Dogma" H. P. Grice and P. F. Strawson (Grice and Strawson 1956) raise several influential objections to Quine's reasoning in "Two Dogmas". One of their objections presupposes that

> Quine requires of a satisfactory explanation of an expression that it should take the form of the pretty strict definition but should not make use of any member of a group of inter-definable terms to which the expression belongs. (Grice and Strawson 1956, p. 234)

Grice and Strawson object that this requirement is in general too strong. In particular, in the case of "analytic" they believe we should be permitted to make use of other terms in the same family, including "synonymous", and "necessary".

This objection rests on a misunderstanding. Quine does not in general require of a definition that it "not make any use of any member of a group of inter-definable terms to which the expression belongs," as Grice and Strawson claim. Quine assumes, instead, that the ordinary language terms "analytic", "synonymous", and "necessary" are not clear enough to help us to settle a boundary between the analytic and the synthetic sentences of the sort Carnap and other scientific philosophers would need to carry out their logical empiricist project. As I noted above, Carnap himself emphasizes this point (Carnap 1942, p. 62).

Grice and Strawson also challenge Quine's claim in §6 of "Two Dogmas" that any sentence, no matter how firmly held, may be revised. If this observation is to be

relevant to the claim that a firmly held sentence is analytic, they reason, then it must be understood as the claim that any sentence can be revised without changing its meaning. Grice and Strawson correctly point out, however, that from the fact that any sentence may be revised it does not follow that any sentence may be revised without changing its meaning. They believe that Quine has missed this obvious point, and so one of his central arguments against analyticity is fallacious (Grice and Strawson, p. 243).

While superficially plausible, this influential criticism rests on three related misunderstandings. First, Quine does not deny that some changes in theory bring about changes in meaning. When he wrote "Two Dogmas," he had already developed his well-known view that "the laws of mathematics and logic … are so central, any revision of them is felt to be the adoption of a new conceptual scheme, the imposition of new meanings on old words" (Quine 1950, p. xiv). His point in §6 of "Two Dogmas" that the truths of logic can be revised is not about meaning, but methodology: it amounts to the claim (reminiscent of the view Carnap expresses in the block quote at the end of §2 above) that no statement we now accept, not even a logical truth that we find obvious, is guaranteed to be part of every scientific theory we will later come to accept (Ebbs 2016). Second, Quine held that we do not need to classify the simple truths of logic as analytic to explain why we firmly accept them and regard rejections of them as changes in their meanings. It is enough, he thinks, that we take the simple truths of logic to be so obvious that apparent evidence that another speaker rejects them is evidence that we have badly translated his words (Quine 1960, pp. 59, 66–67). Third, as I stressed above, Quine's central target in "Two Dogmas" is Carnap's proposed account of the analytic-synthetic distinction. According to Carnap, however, the truth-values of an L-valid (analytic) or a P-valid sentence do not change; a decision to cease affirming such sentences is a decision to cease using the language system of which they are a part and to adopt a language system with different rules. To draw an analytic-synthetic distinction of the sort that Carnap and other scientific philosophers sought, it is therefore not sufficient to show that there are some truths one cannot cease to affirm without ceasing to use the language system of which they are a part, for this is also true of P-valid sentences, as Carnap points out in his reply to "Two Dogmas" (Carnap 1963c, p. 921).

5. Hilary Putnam's rejection of Carnap's analytic-synthetic distinction

Unlike Grice and Strawson, in the early part of his career Hilary Putnam was deeply sympathetic with the principles of scientific philosophy, knew the literature in logical empiricist philosophy of science, and contributed to it himself. Putnam studied with Quine at Harvard, where he did his first year of graduate studies. He then transferred to UCLA, where he studied with Hans Reichenbach. Soon after completing his PhD at UCLA, Putnam also worked closely for a time with Rudolph Carnap. After learning of Quine's criticisms of the analytic-synthetic distinction in 1950 from Carl Hempel (Putnam 2015, pp. 16–17), Putnam became one of the earliest converts to Quine's project of developing a new scientific philosophy that does not rest on the logical empiricists' analytic-synthetic distinction.

In his first major contribution to the debate, "The Analytic and the Synthetic" (Putnam 1962a), Putnam emphasizes that there are some statements in natural language, such as "Bachelors are unmarried," for which (first) there is only one criterion for applying the subject term, in this case "Bachelor", and (second) by this criterion, the sentence is true. The existence of such statements apparently demonstrates that there are some analytic sentences in natural language, contrary to what Quine claims. As Putnam knows, however, this is not a deep challenge to Quine's arguments in "Two Dogmas", for two main reasons. First, as noted above, in §2 of "Two Dogmas", Quine grants that explicit acts of definitional abbreviation create synonymies. The relationship between a word we introduce by an explicit act of definitional abbreviation and the expression we introduce the word to abbreviate is similar to the relationship between our uses of a one-criterion word of an unregimented natural language and the longer phrase of that language that states the generally accepted criterion for applying the word. It is therefore not a big step for Quine to acknowledge the existence in natural language of one-criterion words, such as "Bachelor," and the corresponding sentences, such as "Bachelors are unmarried," to which everyone assents. Quine took this step, acknowledging Putnam, in chapter two of *Word and Object* (Quine 1960, pp. 56–57).

Second, and more important, like Quine, Putnam saw that for any two words at least one of which is not a one-criterion word, the question whether the words are synonymous is at best unclear. In "The Analytic and the Synthetic" Putnam develops and extends this part of Quine's criticism of the logical empiricists' analytic-synthetic distinction by highlighting a range of examples of theoretical statements that are not fruitfully classified as either analytic or synthetic. For instance, before the development of relativity theory, Putnam explains, physicists were unable to see any way in which '$e = ½ mv^2$', an equation for kinetic energy, could be false. They held it immune from disconfirmation by new empirical evidence, and it was reasonable for them to do so. By Carnap's logical empiricist principles, Putnam notes, the methodological role of the equation is best explained by describing it as true by definition of kinetic energy. After Einstein developed relativity theory, however, scientists revised '$e = ½ mv^2$', replacing it with a more complicated equation that fits the new theory, and concluded that '$e = ½ mv^2$', while approximately true, was strictly speaking false, hence not true by definition.

To make sense of such cases, Putnam introduces the idea of a "law-cluster" term, which figures in many different laws of a theory. He observes that we can give up one of the laws in which such a term figures without concluding that the reference of the term has changed. For instance, we can continue to use a given term to refer to kinetic energy while radically changing our theory of kinetic energy. Such terms are, in a word, *trans-theoretical*.

In "It Ain't Necessarily So" (Putnam 1962b) Putnam observes that our theories of the geometry of physical space have changed since the eighteenth century, when the principles of Euclidean geometry were so fundamental to our way of thinking about physical space that we could not then conceive of any alternatives to those principles. Putnam argues that while our theory of physical space has changed radically since the eighteenth century, it is nevertheless correct to regard the terms that scientists in the

eighteenth century used to refer to paths through physical space as trans-theoretical and to conclude that many of the sentences about physical space that scientists accepted in the eighteenth century, such as "The sum of the interior angles of any triangle formed by joining three points in physical space by the shortest paths between them is 180°," are false. Putnam concludes that some statements are so basic for us at a given time that it would not be reasonable to give them up at that time, even if our failure to be able to conceive of alternatives to them is no guarantee that they are true. This observation simultaneously discredits both Kant's view that the theory that physical space is Euclidean is synthetic a priori and the logical empiricists' alternative view that the development of relativistic non-Euclidean theories of space changed the meaning of such theoretical terms as "straight line" and "shortest path between two points in physical space".

Putnam's emphasis on transtheoretical terms is in tension with Quine's thesis that translation between theories is indeterminate (Quine 1960, chapter 2). Each in his own way, however, Putnam and Quine accept that for any given sentence that we now accept, there is no methodological guarantee that we will not one day regard the sentence as false, while still translating our past uses of the words in the sentence into our new theory homophonically. Moreover, like Quine, Putnam rejects the traditional absolute conception of apriority and replaces it with theory-relative methodological observations about which statements we are least likely to revise in the face of unexpected experiences (Putnam 1962a; Putnam 1979).

6. Putnam's externalist semantics and analyticity

Putnam's compelling observations about theory change, first published in the 1960s, discredited the then standard theories of reference and meaning. His proposal that we view some of our terms as law-cluster (i.e. trans-theoretical) terms was a first step away from standard theories. A second step was to extend his notion of trans-theoretical terms, which he first introduced primarily to make sense of cases in which a single inquirer *changes* her view from one time to another, to cases in which two or more inquirers (or speakers) *simultaneously* use a term with the same reference despite large differences in the theories or beliefs they associate with the term. In this key step Putnam observes that we typically assume that ordinary English speakers can use the term 'elm' to refer to elm trees, and 'beech' to refer to beech trees, even if they know very little about elms and beeches, and cannot tell them apart. To distinguish elms from beeches, or to learn about these trees, such ordinary speakers rely on others who know more about them. We rely, in short, on what Putnam calls the division of linguistic labor. He proposes that we reject any theory of reference that implies that ordinary speakers cannot refer to (or think about) elms when they use the term 'elm', even though they do not know much, if anything, about elms, except, perhaps, that they are trees.

Putnam also argues that the references of such "natural kind" terms are dependent in part on the environment in which they are applied, even if it takes years of inquiry and theorizing to discover what the references are. He argues that to discover the reference of a term and learn about its properties is also to clarify what it is true of, and thereby also to

clarify one key component of the meaning of the term, namely, its contribution to the truth conditions of sentences in which it occurs. Appropriating the causal picture of reference that Saul Kripke sketches in *Naming and Necessity* (Kripke 1980) Putnam also theorizes that the meanings and references of a speaker's words are determined in part by causal relations the speaker bears to other speakers in her community and to the environment in which she applies the terms. All these points lead him to his famous conclusion that "meanings ain't in the head" (Putnam 1975a, p. 227). The resulting externalist semantic theory extends and clarifies points that Putnam had already made in "It Ain't Necessarily So," including that many of the statements traditional philosophers had regarded as analytic, such as "Gold is a yellow metal" and "Cats are animals", are, in fact, synthetic and theoretical, and may therefore be undermined by new evidence.

7. Kripke's return to apriority, necessity, and analyticity

Quine's and Putnam's revolutionary work on the methods of inquiry freed Anglo-American philosophy from the narrow confines of Carnap's logical empiricism. Their liberating work created a new intellectual environment in which all types of philosophy, even profoundly unscientific ones, could once again seem worth taking seriously. In the 1970s Saul Kripke took advantage of this opening to develop a possible-world semantics for proper names and kind terms that is rooted not in our reliance on transtheoretical terms in scientific inquiry, but in our everyday intuitions about epistemic and metaphysical possibility (Kripke 1980, pp. 41–42). In a fundamental departure from the spirit of scientific philosophy that runs through the logical empiricists, Carnap, Quine, and Putnam, Kripke's externalist possible-world semantics revives and distinguishes between traditional philosophical concepts of necessity, apriority, and analyticity. Philosophers who applaud Kripke's revival of these traditional philosophical concepts tend to dismiss Carnap, Quine, and Putnam's alternative methodological principles, and to favor methods of inquiry that are very different from those of Carnap, Quine, and Putnam.

Kripke regards the concept of analyticity not as a replacement for and clarification of the traditional concepts of necessity and apriority, as the logical empiricists did, but as less fundamental than, and hence to be defined in terms of, these traditional concepts. He stipulates that an "analytic statement is, in some sense, true by virtue of its meaning and true in all possible worlds [i.e. necessary] by virtue of its meaning.…[so that] something which is analytically true will be both necessary and *a priori*." (Kripke 1980, p. 39) He does not explain in detail what "true in virtue of meaning" comes to, however, and his judgments on particular cases depart from the logical empiricists'. Kripke claims, for instance, that Goldbach's conjecture (that every even number greater than two is the sum of two prime numbers) if true, is necessarily true, but that it is "nontrivial" to claim that one can determine a priori whether or not Goldbach's conjecture is true (Kripke 1980, p. 37). It follows that for Kripke it is nontrivial to claim that Goldbach's conjecture, if true, is analytic. In sharp contrast, Carnap would not rely on the vague traditional term "a priori" to address this question. In Carnap's view, Goldbach's conjecture either is or is not a logical consequence in a given language system *LS* of the empty set of premises; if it is such a consequence, it is analytic in *LS*, otherwise not.

8. Theories of apriority and analyticity after Kripke

To clarify the nature and limits of our supposed a priori knowledge in a way that incorporates something like Kripke's externalist semantics for proper names and kind terms, David Chalmers (Chalmers 1996) and Frank Jackson (Jackson 1998) distinguish between a part of a word's meaning that we can know a priori (its "primary intension") and a part of a word's meaning that we can know only via empirical inquiry (its "secondary intension"). I shall focus here on Chalmers's version of this view, known as two-dimensionalism. Chalmers defines the primary intension of a word as a special sort of function from (agent-centered) worlds to extensions: in a given (agent-centered) word *w*, the primary intension of a word picks out what the extension of the word would be if *w* turned out to be actual (Chalmers 1996, p. 57). To grasp the primary intension of 'water', for instance, we must grasp a function that yields the set of all and only portions of water as value if the actual (agent-centered) world has water in its rivers, lakes, and oceans, and that yields the set of all and only portions of twin water (a superficially similar liquid whose molecular structure is radically unlike H_2O) as value if the actual (agent-centered) world has twin water in its rivers, lakes, and oceans. According to Chalmers, the existence of primary intensions for our words makes it possible for us to identify items to which our words apply, and to make empirical inquiries into the nature of those items. In this way we may come to believe that water is H_2O, just as Kripke and Putnam claim. Such a discovery yields a secondary intension expressed by our actual applications of 'water'. If the stuff to which we actually apply our word 'water' is H_2O, the secondary intention of our actual applications of 'water' picks out the H_2O in every counterfactual world (Chalmers 1996, p. 57).

As noted above, the semantic externalism developed by Putnam and Kripke implies that many of the statements traditional philosophers had regarded as analytic, such as "Gold is a yellow metal" and "Cats are animals", are, in fact, synthetic and theoretical, and may therefore be undermined by new evidence. Chalmers concedes the point, but insists that our knowledge of the primary intensions of our words, including "gold" and "cat", is a priori. For instance, he claims we can know a priori what *would be the case* if we *were* to discover that the things to which we apply our word 'cat' are not animals. As Putnam stresses, however, statements that we cannot now imagine giving up without changing the subject are not thereby guaranteed to be true. What we actually judge when we find ourselves in a previously imagined situation almost always trumps our earlier speculations about what we *would* say if we *were* to find ourselves in that situation. Chalmers concedes this point also, but retreats to the view that one might always be mistaken about one's a priori claims, so the fact that we have revised many of our previously deeply held beliefs about how to apply a given word does not show that we cannot in principle know a priori how to apply that word in any given possible world. The methodological significance of Chalmers's conception of the a priori is unclear, however, since in practice, on his view, we may be no more confident of our a priori claims about how to apply our terms than we are of many of our empirical theories, and for the same basic reason—new evidence may undermine them.

Paul Boghossian aims to do better than this for at least some of the propositions that we take ourselves to know a priori. He argues that some of the sentences we accept implicitly define the meanings of our logical constants and that these sentences are epistemically analytic for us, in the sense that our grasp of their meanings constitutes epistemological warrant for believing the propositions they express, and our warrant for believing the propositions they express is not defeasible by any future empirical evidence (Boghossian 1996, p. 362). In sharp contrast to Carnap, who explicates analyticity solely in terms of *explicitly* adopted rules, including rules that implicitly define some of the words of a language, Boghossian proposes that we can define analyticity in terms of *implicitly* adopted implicit definitions of words. A central problem for his position is that he does not provide an informative criterion for distinguishing meaning constituting sentences from other sentences we tacitly accept now but might revise later. Unless and until we know such a criterion, or we know at least that we judge in accord with it, it will not be transparent to us which of our sentences, if any, is analytic in Boghossian's sense. A related and even more fundamental problem for Boghossian's position is that the quinean methodological considerations discussed above apparently show that "it is objectively indeterminate which principles are true by virtue of meaning and which are substantive" (Harman 1996, p. 397).

One might respond to this quinean methodological challenge by claiming that a statement that members of a community take to be true and indefeasible should not be translated by a false, defeasible statement. On this view, as developed by Cory Juhl and Eric Loomis, if s is a sentence that a speaker stipulates to be true and indefeasibe, then s' is a good translation of s only if s' is also true and indefeasible (Juhl and Loomis, pp. 224–225). Quine and Putnam reject this view of translation: when we change our theory, they argue, we sometimes also thereby change our view of how best to translate statements we or others previously regarded as true and indefeasible. This theory-relative account of translation is supported by cases from the history of science, such as Putnam's Euclidean space case, but there is no completely neutral argument for or against it.

A very different tack is to abandon the search for a theory of epistemic analyticity and focus instead on explicating the idea that analytic truths are "true in virtue of meaning". Following Gillian Russell, for instance, one might regard the sentence "I am here now" as analytic, in the sense that any utterance of it is guaranteed to be true, yet also contingent, since wherever a person is when she utters it at a given time, she might have been elsewhere at that time. Generalizing from such examples, Russell isolates a component of word's meaning that she calls its "reference determiner," and argues that analyticity, or truth in virtue of meaning, should be understood as truth in virtue of reference determiner. A consequence of her view is that analyticity cannot play the epistemic role that some philosophers have imaged for it, since, while "analytic justification is justification on the basis of reference determiners alone," competent speakers might not know what the reference determiners of their words are, and therefore might not be able "to avail themselves of analytic justifications" (Russell 2008, p. 176).

9. Analyticity and indeterminacy of meaning

Whether one finds a given explication of analyticity acceptable depends in part on whether one believes that meaning is determinate in the ways that the account requires. Some philosophers have argued that it is incoherent to suppose that meaning is indeterminate, and hence that criticisms of explications of analyticity that imply that meaning is indeterminate are unacceptable for purely logical reasons, independent of empirical inquiry. According to Grice and Strawson, for instance, "If talk of sentence-synonymy is meaningless, then … talk of sentences having a meaning at all must be meaningless too" (Grice and Strawson 1956, p. 146). As Quine points out, however, baring philosophical prejudices to the contrary, it is neither logically incoherent nor puzzling to reject this conditional (Quine 1960, p. 206). For similar reasons, John Searle's efforts (in Searle 1987) to reduce Quine's thesis of the indeterminacy of reference to absurdity are also ultimately unsuccessful (Ebbs 2011, pp. 624–627). There is unlikely to be a decisive demonstration that indeterminacy of meaning is incoherent.

In the 1960s and 70s, at the start of the cognitive revolution in linguistics and psychology, some prominent linguists and philosophers claimed that empirical research in these sciences, especially Noam Chomsky's work and his criticisms of Quine's indeterminacy thesis (Chomsky 1969), establishes that meaning is determinate, despite Quine's and Putnam's methodological arguments to the contrary. It is still often claimed, for instance, that Quine's arguments for indeterminacy of meaning rest on an outdated and discredited linguistic behaviorism (Burge 2010, pp. 223–227; Anthony 2016, pp. 26–29). But this judgment is not universally shared by even by those who oppose Quine's arguments against analyticity and the determinacy of meaning. According to Jerrold Katz, who rejects Quine's arguments for indeterminacy, for example, "Quine's behaviorism merely takes linguists out of their armchairs and puts them in the field facing the task of having to arrive at a theory of language on the basis of the overt behavior of its speakers in overt circumstances" (Katz 1990, p. 180; quoted in Hylton 2007, p. 374 note 8). The feeling nevertheless persists that cognitive psychology and linguistics somehow demonstrate that Quine and Putnam are wrong about meaning and analyticity. Eliot Sober writes, for instance, that, "If cognitive science produces predictive and explanatory theories that attribute to speakers a knowledge of what their terms mean, then the concept of meaning is on a safe footing, and so is the concept of analyticity" (Sober 2000, p. 277; see also Williamson 2007, p. 50). Unless we are confident that successful explanatory theories in cognitive science and linguistics can support philosophical claims about the analyticity of certain public language sentences, however, we should hesitate to affirm Sober's conditional. And doubts about the scientific respectability of contemporary philosophical theorizing about meaning and representation have been growing. Chomsky himself has apparently changed his mind on this crucial point, arguing in 2000 against standard truth-conditional semantics that when it comes to reference, the proper scientific focus is "internal" and "syntactical", not semantical in a sense that relates expressions to non-linguistic things (Chomsky 2000, p. 42). Despite Chomky's doubts, a large group of philosophers and linguists continue to develop externalist semantics for natural languages and to assume that the semantic relations they describe are objective and determinate. Among cognitive scientists more generally,

however, as William Ramsey reports, there is "disarray and uncertainty" about the very ideas of meaning and mental representation, including "disagreements about how we should think about mental representation, about why representations are important for psychological and neurological processes, about what they are supposed to do in a physical system, about how they get their intentional content, and even about whether or not they actually exist" (Ramsey 2007, p. xi). In short, the proliferation over the past 60 years of competing accounts of the nature and role of meaning and mental representation in philosophy, linguistics, and cognitive psychology raises serious doubts about the often-repeated claim that cognitive science establishes that "the concept of meaning is on a safe footing, and so is the concept of analyticity" (Sober 2000, p. 277). Such doubts call for an open-minded investigation that starts by getting clear, as I try to do in outline above, on the strengths and weakness of the most promising theories of analyticity.

Acknowledgements

For comments on previous drafts I thank Kate Abramson, Albert Casullo, Cory Juhl, and Savannah Pearlman. Parts of §§5–6 are lifted verbatim from my paper "Putnam on Methods of Inquiry" (*Argumenta* 3, December 2016, pp. 157–161), and the first two paragraphs of §8 include material from §§6.9–6.10 of my book *Truth and Words* (Oxford, 2009).

Works Cited

Anthony, Louise, 2016. "Things Oughta Make Sense," *Proceedings and Addresses of the American Philosophical Association*, Volume 90 (November 2016), pp. 21–39.

Auxier, R., Anderson, D. R., and Hahn, L. E. 2015. *The Philosophy of Hilary Putnam*. Chicago, Illinois: Open Court.

Boghossian, Paul, 1996. "Analyticity Reconsidered," *Noûs*, 30:3: 360–91.

Burge, Tyler, 2010. *The Origins of Objectivity*. Oxford: Oxford University Press.

Carnap, Rudolf, 1934. "The Task of the Logic of Science," an English translation of a paper originally published in German in 1934. In McGuinness, Brian, (ed.), *Unified Science: the Vienna Circle Monographs Series Originally Edited by Otto Neurath*. Dordrecht: D. Reidel, 1987, pp. 46–66.

1936–37. "Testability and Meaning," Parts I and II, from *Philosophy of Science*, Vol. III, 1936, pp. 419–471, and Vol. IV, 1937, pp. 1–40.

1937. *The Logical Syntax of Language*. An expanded version of what was originally published in German in 1934. Translated by A. Smeaton. London: Kegan Paul, Trench, Trubner & Co.

1942. *Introduction to Semantics.* Cambridge, Mass.: Harvard University Press.

1956. "The Methodological Character of Theoretical Concepts," in Herbert Feigl and Michael Scriven, eds., *The Foundations of Science and the Concepts of Psychology and Psychoanalysis. Minnesota Studies in the Philosophy of Science, Volume 1* (University of Minnesota Press), pp. 38-76.

1963a. "Intellectual Autobiography," in Schilpp 1963, pp. 3–84.

1963b. "Carl G. Hempel on Scientific Theories," in Schilpp 1963, pp. 958–966.

1963c. "W. V. Quine on Logical Truth," in Schilpp 1963, pp. 915–922.
1966. Martin Gardner, ed., *The Philosophical Foundations of Physics: An Introduction to the Philosophy of Science*. New York: Basic Books. Reprinted in a slightly revised edition as *An Introduction to the Philosophy of Science*, by Dover in 1995. References in the text are to the Dover edition.
Chalmers, David, 1996. *The Conscious Mind*. Oxford: Oxford University Press.
Chomsky, Noam, 1969. "Quine's Empirical Assumptions," in Davidson, D., and Hintikka, J., eds., *Words and Objections* (Dordrecht: D. Reidel, 1969), pp. 53–68.
2000. "Explaining Language Use," in Noam Chomsky, *New Horizons in the Study of Language and Mind* (Cambridge: Cambridge University Press, 2000), pp. 19–45.
Domski, Mary, and Dickson, Michael, 2010. *Discourse on a New Method*. Chicago, Illinois: Open Court.
Ebbs, Gary, 2011a. "Carnap and Quine on Truth by Convention," *Mind* 102/478: 193–237; reprinted in Ebbs 2017a, pp. 57–94.
2011b. "Quine Gets the Last Word," *The Journal of Philosophy* 108/11: 617–632; reprinted in Ebbs 2017a, pp. 113–127.
2015. "Reading Quine's Claim that No Statement is Immune to Revision," in Janssen-Lauret and Kemp 2016, pp. 123–145; reprinted in Ebbs 2017a, pp. 168–187.
2017a. *Carnap, Quine, and Putnam on Methods of Inquiry*. Cambridge: Cambridge University Press.
2017b. "Reading Quine's Claim that Definitional Abbreviations Create Synonymies," in Ebbs 2017a, pp. 128–143.
Friedman, Michael 2010, "Synthetic History Reconsidered," in Domski and Dickson 2010, pp. 573–813.
Frost-Arnold, Greg. 2013. *Carnap, Tarski, and Quine at Harvard: Conversations on Logic, Mathematics, and Science*. Chicago, Illinois: Open Court.
Gödel, Kurt. 1931. "On Formally Undecidable Propositions of *Principia Mathematica* and Related Systems I", in Solomon Feferman, John W. Dawson, Jr., Stephen C. Kleene, Gregory Moore, Robert N. Solovay, and Jean van Heijenoort (eds.), *Kurt Gödel, Collected Works*, Vol. I (Oxford: Clarendon Press) pp. 144–195.
Grice, H.P. and Strawson, P.F., 1956. "In Defense of a Dogma," *The Philosophical Review*, Vol. 65, No. 2: 141–158.
Harman, Gilbert, 1996. "Analyticity Regained?" *Noûs* 30: pp. 392–400.
Hempel, Carl G., 1963. "Implications of Carnap's Work for the Philosophy of Science," in Schilpp 1963, pp. 685–709.
Hookway, Christopher, 1988. *Quine*. Stanford: Stanford University Press.
Hylton, Peter, 2007. *Quine*. New York: Routledge.
Jackson, Frank, 1998. *From Metaphysics to Ethics: A Defense of Conceptual Analysis*. Oxford: Oxford University Press.
Janssen-Lauret, Frederique, and Kemp, Gary, eds., 2016. *Quine and his Place in History*. New York: Palgrave Macmillan.
Juhl, Cory, and Loomis, Eric, 2010. *Analyticity*. New York: Routledge.
Kant, Immaneul, 1871. *Critique of Pure Reason*. Originally published in German. Translated by Norman Kemp Smith. New York: Macmillan Press, 1929.

Katz, Jerrold, 1990. "The Refutation of Indeterminacy," reprinted in R. Barrett and R. Gibson, eds., *Perspectives on Quine* (Oxford: Blackwell, 1990), pp. 177–197.
Kripke, Saul 1980. *Naming and Necessity*. Cambridge, Mass.: Harvard University Press.
Putnam, Hilary 1962a. "The Analytic and the Synthetic," reprinted in Putnam 1975c, pp. 33–69.
1962b. "It Ain't Necessarily So," reprinted in Putnam 1975b, pp. 237–249.
1975a. "The Meaning of 'Meaning'," reprinted in Putnam 1975c, pp. 215–271.
1975b. *Mathematics, Matter, and Method: Philosophical Papers, Vol. 1*. Cambridge: Cambridge University Press.
1975c. *Mind, Language, and Reality*: *Philosophical Papers, Volume 2*. Cambridge: Cambridge University Press.
1979. 'Analyticity and Apriority: beyond Wittgenstein and Quine', first published in P. French et al, (eds.), *Midwest Studies in Philosophy*, Volume IV (Minneapolis: University of Minnesota Press); reprinted in Putnam 1983c, pp. 115–138.
2015. "Intellectual Autobiography," in Auxier, Anderson, and Hahn 2015, pp. 3–110.
Quine, W. V. 1936. "Truth by Convention," first published in O. H. Lee, (ed.), *Philosophical Essays for A. N. Whitehead*, New York, NY: Longmans, 1936; reprinted in Quine 1976, pp. 77–106.
1953a. "Two Dogmas of Empiricism," in Quine 1953b, pp. 20–46.
1953b: *From a Logical Point of View*. Cambridge, Mass.: Harvard University Press.
1960: *Word and Object*. Cambridge, Mass.: MIT Press.
1963. "Carnap and Logical Truth," in Schilpp 1963, pp. 385–406.
1971. "Homage to Carnap." First published in Roger C. Buck and Robert S. Cohen, (eds.), *Boston Studies in the Philosophy of Science*, Vol. 8 (Dordrecht: D. Reidel Publishing Co., 1971); reprinted in Richard Creath, ed., *Dear Carnap, Dear Van* (Berkeley: University of California Press, 1990), pp. 463–66.
1994. "Exchange between Donald Davidson and W. V. Quine following Davidson's lecture," *Theoria* 60: pp. 226–231.
1986a: "Intellectual Autobiography," in Hahn, Lewis E. and Paul A. Schilpp (eds) 1986: *The Philosophy of W. V.* Quine (La Salle, IL: Open Court), pp. 3–46.
Ramsey, William M. 2007. *Representation Reconsidered*. Cambridge: Cambridge University Press.
Russell, Gillian 2008. *Truth in Virtue of Meaning*. Oxford: Oxford University Press.
Schilpp, Paul A. (ed.), 1963. *The Philosophy of Rudolf Carnap*. La Salle, IL: Open Court.
Searle, John R., 1987, "Indeterminacy, Empiricism, and the First Person," *The Journal of Philosophy*, Vol. 84, No. 3: 123–46.
Soames, Scott. 2003. *Philosophical Analysis in the Twentieth Century, Volume 1: The Dawn of Analysis*. Princeton, New Jersey: Princeton University Press.
Sober, Elliot, 2000. "Quine," *Proceedings of the Aristotelian Society*, Supplementary Volume 74: 237–280.
Tarski, Alfred. 1936. "The Concept of Truth in Formalized Languages," in Alfred Tarski, *Logic, Semantics, Meta-Mathematics: Papers from 1923 to 1938*, 2nd ed., translated by J. H. Woodger, edited by John Corcoran (Indianapolis: Hackett, 1983), pp. 152–277.
Williamson, Timothy 2007: *The Philosophy of Philosophy*. Oxford: Basil Blackwell.