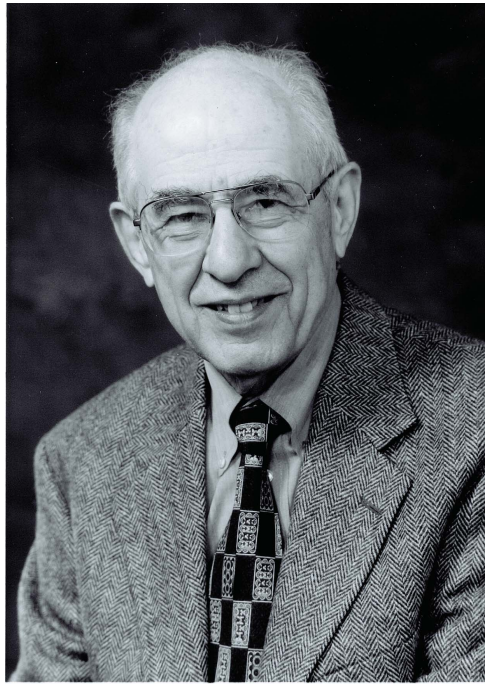
 Mind & Matter Vol. 14(1), pp. 87–95

Hilary Putnam (1926–2016): A Lifetime Quest to Understand the Relationship between Mind, Language, and Reality

David Leech Anderson
Department of Philosophy
Illinois State University, Normal, USA



Hilary Putnam was one of the most influential philosophers of the 20th century. His peers have called him "one of the finest minds I've ever encountered" (Noam Chomsky) and "one of the greatest philosophers this nation has ever produced" (Martha Nussbaum). Standard obituaries covering his "life and times" are available in the mass media.[1] This essay

[1]See, e.g., the New York Times (`www.nytimes.com/2016/03/18/arts/hilary-putnam-giant-of-modern-philosophy-dies-at-89.html`), the Guardian (`www.theguardian.com/books/2016/mar/14/hilary-putnam-obituary`), or the Huffington

will focus exclusively on the legacy he has left from his published works and from the way he has modeled a philosopher's life.

Hilary Putnam was a towering figure in philosophy of language, philosophy of mind, and in metaphysics, and special attention will be paid (below) to those contributions. His influence, however, extends well beyond those boundaries. He has published works in dozens of different research areas, of which the following is only a small sample. Early in his career he contributed (with Martin Davis, Yuri Matiyasevich, and Julia Robinson) to the solution of Hilbert's 10th problem, a major result in mathematics which, together with other related accomplishments, would have warranted an appointment in a mathematics department. While his central focus was not ethics and political philosophy he advanced a sustained attack on the fact-value distinction – a central presupposition of logical positivism and a position that he thought continued to influence research in many areas to deleterious effect. His work in this area was applied not only to ethics and the philosophy of science but in a recent book (edited with Vivian Walsh) he argued that in economic theory (as elsewhere) facts and values are so "entangled" that there can be no such thing as value-free economic theory.[2] In the second half of his life, he turned in earnest to his Jewish heritage, learned Hebrew, was bar mitzvahed at age 68, and wrote about Jewish philosophy.[3]

A list of the many and diverse contributions he has made to our intellectual life would be too lengthy for this venue and, in the form of a list, would not adequately capture the unifying elements of his philosophical work. As important as are the theories and the arguments he has left us, of equal importance is the ideal of philosophical inquiry that he has embodied throughout his career. He is a rare example of someone who relentlessly pursued an answer to one central question, never wavering as he sought richer, more nuanced answers to the question: "What is the nature of the mind-world relation?"

Throughout his career, and in virtually every one of his obituaries, great emphasis has been placed on his uncommon penchant for "changing his mind" – for attacking philosophical theories that he himself created. The temptation is to attribute this to a quirky personality trait, as if he had commitment issues of some kind. That, of course, is to completely misunderstand not only the man, but the nature of his philosophical contribution. To better understand the legacy he leaves us, this essay considers the work he has pursued on the basic question of the mind-world

Post (`www.huffingtonpost.com/martha-c-nussbaum/hilary-putnam-1926-2016_b_9457774.html`).

[2] H. Putnam and V. Walsh, eds. (2012): *The End of Value-Free Economics*, Routledge, New York.

[3] H. Putnam (2009): *Jewish Philosophy as a Guide to Life: Rosenzweig, Buber, Levinas, Wittgenstein*, Indiana University Press, Bloomington.

relation.

On the one hand, it seems perfectly obvious that there exists an objective, physical world and it is easy to theorize as if we have an omniscient, God's eye access to its nature. But we don't. We are finite creatures, with a very particular (and contingent) epistemic access to that reality. Our sensory and cognitive access to reality is limited by the constraints and particularities of this biological organism. So we ask:

1. What is the nature of cognition? What are minds that they are capable of thinking about distant galaxies to which we lack physical access?
2. What is language? How can linguistic symbols serve as the medium which carries cognitive content – and how do they hook onto the physical world?
3. And what of reality itself? Since we can't assume that our cognitive access to reality is literally transparent, as of an omniscient mind, how might the world that we inhabit be conditioned by the cognitive and sensory faculties which mediate our access to it?

Putnam has made countless contributions to the answers that we are now able to give to these questions. He insisted that the three issues were inextricably entwined, or (as he was fond of saying) the three "take in each other's laundry". Just as Dante needed a guide (like Virgil) to direct his steps through the underworld, consider Hilary Putnam our guide as we navigate the perplexities surrounding mind, language, and reality. By reviewing his legacy in these areas of inquiry, we will see much of the progress that has been made on these questions during the past 75 years. Putnam has been at the center of the most revolutionary developments.

While a student of logical positivists (including Hans Reichenbach and Rudolf Carnap), Putnam was unflinching in his attacks on the excesses of that philosophical movement. Logical positivism offered a very simplistic account of the relationship between thought/language and the physical world, which emphasized empiricism (and human sensory experience as reported in "observation sentences") at the expense of common sense realism about objects in the world. For some positivists, this took the extreme form of phenomenalism, the view that physical objects are nothing more than "permanent possibilities of experience". For Rudolf Carnap, one of Putnam's teachers, it resulted in a rejection of a realist interpretation of the theoretical entities mentioned in scientific theories. Such terms ("electron", "microbe") were merely of instrumental value in predicting future "observation statements" and so were not translated with ontological import. This revealed a tension within 20th century analytic philosophy between a methodological commitment to empiricism and an ontological commitment to naturalism. Putnam judged that logical positivism

emphasized empiricism at the expense of realism and he advanced now famous arguments for both scientific realism – arguing that the success of science would be a "miracle" if the entities referred to did not really exist,[4] and that mathematical realism – the existence of numbers – is "indispensable" to scientific theorizing.[5]

Putnam also took on the two theories of mind that dominated during the 1950s: behaviorism and mind-brain identity. Against behaviorism he offered a thought experiment about a community of "super-spartans" who experience (subjective) pain but never evince behavior expressive of that pain[6] and against the identity thesis he offered a new theory of mind, functionalism, that has dominated cognitive science and philosophy ever since.[7] Drawing inspiration from his own work in the theory of recursive functions, Putnam proposed that the essence of mental states is not what they are made of but how they function. While today functionalism is usually treated as a reductive theory (compared to several contemporary alternatives), at the time it was a rejection of very reductive alternatives in favor of a theory that was completely catholic with respect to the metaphysical stuff of which minds might be made. Whether a human organism, alien organism, silicon-based machine, or immaterial soul, each would have the same type of mental state (e.g., belief or pain) so long as they shared the same function (implemented the same software program) within the intelligent system. His original "machine functionalism" gave way to a more flexible "causal role functionalism" which defines mental states in terms of the causal role it plays with regard to its intake of information from the environment, its causal relations to other mental states, and its contribution to system behavior. For over forty years, functionalism has been the dominant theory of mind.

Putnam famously rejected functionalism[8] because it couldn't capture the richness of the phenomena. He argued that plasticity of mental states between different cognitive agents and even between states of the same agent at different times made possible sameness of mental state in the absence of sameness of functional properties. Yet, at the end of his life, after many twists and turns in the development of his thinking, he still described his position as a form of "liberal functionalism" – insisting that

[4]H. Putnam (1962): What theories are not. In *Logic, Methodology and Philosophy of Science*, ed. by E. Nagel, P. Suppes, and A. Tarski, Stanford University Press, Stanford, pp. 240–251.

[5]H. Putnam (1975): What is mathematical truth? In *Mathematics, Matter, and Method*, Cambridge University Press, Cambridge, pp. 60–78.

[6]H. Putnam (1963): Brains and behavior. In *Analytical Philosophy Vol. 2*, ed. by R.J. Butler, Blackwell, Oxford, pp. 1–19.

[7]H. Putnam (1960): Minds and machines. In *Dimensions of Mind*, ed. by S. Hook, New York University Press, New York, pp. 138–164.

[8]H. Putnam (1988): Why functionalism didn't work. In *Representation and Reality*, MIT Press, Cambridge, pp. 73–89.

there remains an important truth in his original insight.

Next we turn to language. Until the late 1960s, the dominant view of how words come to refer to objects in the world had been little changed for thousands of years. Locke gave modern expression to this standard internalist semantics in his "way of ideas". What words mean, how they pick out the objects to which they refer, is determined by the "ideas" (the properties) that thinkers consciously "associate" with those terms. The properties (odorless, colorless, tasteless, thirst-quenching liquid) that I associate with a particular term (such as water) just are the meaning of the word and the word refers to any object that has those properties. This definite description theory of reference seemed not only the best theory, but the only conceivable theory. If we say that "water" means "odorless, colorless, tasteless, thirst-quenching liquid", how could we all possibly be wrong about the meaning of our own language?

Well, today, most philosophers and cognitive scientists think that we were all wrong, due primarily to the arguments of Putnam and Saul Kripke who, working independently, were reconceiving of how language works, insisting that "the world" – one of our two poles – had not been given its due in fixing the meaning of our words. Putnam began teaching these ideas in 1966, but they got little traction until he finished a paper at the end of 1972 titled "The Meaning of 'Meaning'", which had the additional support of Saul Kripke's famous 1970 lectures on "Naming and Necessity".[9] The Kripke-Putnam "new theory of reference" (also called the "causal theory" or "direct theory") claimed that natural kind terms and proper names are not definite descriptions but are rigid designators. Their semantic content is indexically fixed, most commonly by an act of baptism. A baby's name is causally connected to a specific baby at baptism (Kripke) and natural kind terms, like "water", can be causally connected to a particular kind of physical stuff by ostension – "that stuff is called 'water'" (Putnam). Putnam argued that the meaning of the term "water" was not captured by any description (odorless, etc.) but was attached to that very kind of stuff (whatever its deep explanatory properties turn out to be). If our scientists are correct and water turns out to be H_2O, then "water" refers to all and only those things that are H_2O.

Virtually overnight (at least in academic time), a majority of philosophers and cognitive scientists were convinced by these arguments. How could such a thing happen? What kind of argument could they have given to overwhelm 3000 years of philosophical momentum? It is remarkable, but they gave no principled arguments. They just told stories (which philosophers usually call "thought experiments") and then asked the reader "Is the stuff in the story 'water' or not?" The reader's answer

[9] H. Putnam (1975): The Meaning of "Meaning". In *Mind, Language, and Reality*, Cambridge University Press, Cambridge, pp. 215–271. S. Kripke (1972/1980): *Naming and Necessity*, Harvard University Press, Cambridge.

to the question, confirmed the new theory.

Putnam's famous thought experiment that launched the new theory is about Oscar, who lives on Earth and Twin Oscar who inhabits Twin Earth that is in every way exactly like Earth except everywhere there is H_2O on Earth there is XYZ on Twin Earth. Since everything, except the microstructure of water, is the same on both planets the question is: Do all of the words in Twin English (the language Twin-Earthers speak) mean the same thing as they do in English? The obvious difference might be the meaning of the term "water". If Locke is correct, however, and "water" just means "odorless, colorless, tasteless, thirst-quenching liquid" then the word "water" has the same meaning in both languages. Since H_2O and XYZ are both odorless, colorless, etc., they are both instances of "water" and of twin-"water". Putnam disagrees. He claims that if Oscar were to visit Twin Earth and learn that their lakes do not contain H_2O but a fundamentally different kind of stuff, XYZ, Oscar would say:
"It may look like water but it is not really water because it is not H_2O."
Most readers asking themselves the question "Would I call XYZ 'water'?" answered in the negative. And when they did, the externalist revolution was born. A negative answer to that question refutes Locke' definite description theory and confirms the new theory. One of Putnam's most famous sayings – "meanings just ain't in the head" – became a rallying cry.

The early arguments were just about proper names and natural kind terms, but it soon became a global claim about all mental content. By the time Putnam advanced his famous "brains in a vat" argument,[10] one of the least contentious premises of that argument was the claim that "bearing a causal connection (of the appropriate kind)" was a necessary condition for any word (even descriptive words) to genuinely refer to objects and properties in the physical world. During this period, tracking theories of intentionality (or mental content) broke onto the scene exploiting the Kripke-Putnam insight and offering powerful new theories about how thought and language connects to objects in the world.[11]

If one stopped here in the survey of Putnam's contribution to our understanding of the relationship between mind, language, and reality one would assume that Putnam could be easily pigeon-holed. Clearly, he is committed to the following: (a) *Metaphysical and epistemological realism*: The external world consists in mind-independent objects and we are capable of having knowledge of them; (b) *Semantic externalism*: Words refer to whatever real, mind-independent physical objects those

[10] H. Putnam (1981): Brains in a vat. In *Reason, Truth, and History*, Cambridge University Press, Cambridge, pp. 1–21.

[11] For examples of tracking theories, see Fred Dretske's *Knowledge and the Flow of Information* of 1981 and Ruth Millikan's *Language, Thought, and Other Biological Categories* of 1984.

words reliably track; (c) *Long-armed functionalism*: Mental state types (like beliefs and pains) are individuated by the causal role each plays in the economy of the organism (that is the "functionalist" part) and the content of each mental state is fixed by whatever objects in the external environment it is causally connected to (that is the "long-armed" part).

In 1976, this was Putnam's philosophical profile and his mantel was secure as the leading light advocating for metaphysical and scientific realism. Not only had he offered the most compelling ("miracle") argument for scientific realism, but he had just delivered what many thought was the coups de grace against idealism and antirealism – his externalist semantic theory. The very essence of language is now grounded in causal interactions between real, physical states of the agent and real, physical objects in the external world. Putnam had built a realist juggernaut. And yet, in December of 1976, as president of the Eastern Division of the American Philosophical Association, he shocked the philosophical world in his presidential address titled "Realism and Reason" by announcing that he was now convinced that metaphysical realism is incoherent.

Putnam was a fan of P.F. Stawson's modern reading of Immanuel Kant's critical philosophy (*The Bounds of Sense*, 1966) and became convinced that Kant was right that our cognitive and sensory faculties help to constitute the empirical world that we inhabit they do not simply discover what Putnam called a ready-made world. His former realist comrades in arms were aghast. Michael Devitt, a student of both Putnam and Quine, wrote a paper titled "The Renegade Putnam" making no effort to hide his chagrin at his mentor's reversal. For Putnam, though, it wasn't a betrayal of any fundamental principle. His quest remained the same: To understand the richness of the mind-language-world relationship, doing justice to both poles, "mind" and "world". He became convinced by Michael Dummett's arguments that the only coherent theory of how anyone comes to understand the meaning of a language is by learning the "verification conditions" of its statements – not the realist truth-conditions. He judged that the combination of realism plus functionalism plus semantic externalism pushed the pendulum too far, by overemphasizing the reductive contribution of an objective world.

The most surprising, and original twist that Putnam brought to the realism-antirealism debate was the insight that semantic externalism does not automatically settle the realism-antirealism dispute in favor of realism, but complicates the picture in ways not fully appreciated. It is a basic tenet of realism that radical skepticism is coherent (that it is possible that our beliefs are systematically false). Antirealists traditionally reject this tenet insisting that how things seem (at least ideally and in the long run) cannot be fundamentally different from how they really are – and so radical skepticism is incoherent. Appealing only to the causal requirement on reference and disquotation in one's language, Putnam's "brains in a

vat (BIV) argument" claims that we can know that we are not a brain in a vat, thus ruling out radical skepticism. But if skepticism is incoherent, that is precisely what antirealism affirms and realism denies. Thus, it seems, that semantic externalism leads to a rejection of realism.

Hundreds of articles have been written about the BIV argument and there is no consensus regarding its soundness. But the past thirty-five years have shown since its publication that semantic externalism does indeed have revisionist consequences for how we are to understand our cognitive access to a physical world. Philosophers continue to struggle with how to reconcile the internalist psychological fact that I seem to have immediate access to the content of my own thoughts with the (widely held) conclusion that semantic externalism rules out "self-knowledge" with regard to the content of my own mental states. Putnam has encouraged serious reflection on the consequences of the externalist revolution.

Arguments like the BIV argument and "the model-theoretic argument" (another famous Putnam argument against realism) lead him to defend "internal realism", a view that he eventually admitted was not a form of realism after all, but of antirealism. For almost two decades he challenged the realist orthodoxy of the time, insisting that standard realist versions of long-armed functionalism had unresolved problems. In the 1990s, however, he recanted that position believing that he had swung the pendulum too far back in the direction of the "mind" leaving the "world" too much a product of the mind's operations. He described his next position as a new kind of "naive realism". This coincided with the emergence of two new theories – disjunctivism and representationalism – that Putnam flirted with but ultimately rejected. He was convinced that his semantic externalism and a kind of realism was still a crucial part of the picture, but that many current theories were going too far and reducing everything mental to externalist relations.

To capture what was missing, he turned his energies in the last years of his life to a study of perception, co-authoring several papers with Hilla Jacobson.[12] They proposed a new theory, "transactionism", defending a distinction between "sensation" and "apperception" and insisting that there are unconceptualized perceptions. In the process of this work, Putnam made an unexpected concession to defenders of phenomenal consciousness or qualia that he had never made before. (Qualia are properties of mental states according to which "there is something that it is like" to be in that state, like the feel of a pain or the look of a sunset). He made the concession to Ned Block, another student of his and a defender of a robust kind of qualia that is not reducible to the mere informational content about objects that semantic externalism can deliver. He says:[13]

[12]The last one of these papers is H. Jacobson and H. Putnam (2016): Against perceptual conceptualism. *International Journal of Philosophical Studies* **24**(1), 1–25.

[13]H. Putnam (2012): How to be a sophisticated naive realist. In *Philosophy in*

> ... I agree with Block's claim that qualia are probably identical with brain events/states. Since qualia are what Block calls "mental paint", that means that I do not now see Block's "phenomenism" as incompatible with Jacobson's and my "transactionalism".

Currently, one of the major disputes in the philosophy of mind is that between functional versus phenomenal intentionality. What role (if any) do qualia play in determining the content of mental states and how might they be integrated with functionalism and externalism? As Putnam left us, he was turning his remarkable intellect to that question. We will feel the loss of not having him as our guide through those challenging waters.

an Age of Science, ed. by M. De Caro and D. Macarthur, Harvard University Press, Cambridge, pp. 624–639.