

“Causality-Dependent Consciousness and Consciousness-Dependent Causality”

David Leech Anderson

[Penultimate Draft, Published in *Journal of Consciousness Studies* 19, No. 5–6, 2012, pp. 12–39]

1. Introduction

“Metaphysics” has long been the name for the sub-discipline in philosophy which explores the question “What things ultimately exist in the universe?” and for the things that do exist, “What is their fundamental nature?” The term “metaphysics” has had such a complex and some might say unseemly history that it is now more common to describe these as ontological questions rather than metaphysical ones – although I will use the two terms synonymously. One’s “ontology” consists of the list of everything included in one’s inventory of “the ultimate furniture of the universe,” together with a sufficiently detailed account of what kinds of things the terms on that list refer to.

Both “consciousness” and “causality” have been and continue to be the center of heated ontological disputes, focusing on both their nature and their very existence. Entire courses in philosophy are regularly taught on each of these subjects and dozens of books and articles are published each year on these topics. Of the two, the causality debates go back the furthest, but the recent flurry of activity focusing on consciousness is quickly making up for lost time.

Frequently in philosophy, research into one ontological domain (say, “personal identity”) will inevitably engage issues central to a neighboring ontological domain (say, the nature of “minds”). This is definitely *not* the case with the two topics at hand. While it may be that the two sub-disciplines *ought* to engage one-another, they rarely do. Most volumes dedicated to the nature of causality never even mention ontological controversies over the nature of consciousness, not because the authors fail to know about such controversies (of course, they do) but because they do not judge that the one topic implicates the other topic in any essential way. For those not familiar with the literature, this point can easily be made by appeal to two recent publications that give an overview of the current state of “causality” studies in philosophy. Phil Dowe’s entry on “Causal Processes” in the respected *Stanford Encyclopedia of Philosophy* (Dowe, J. 2009) is an excellent paper, reviewing classic and contemporary contributions to the study of causality. Nowhere in Dowe’s paper does the word “conscious” or “consciousness” even appear. Ned Hall’s paper in *Philosophy Compass*, “Philosophy of causation: blind alleys exposed; promising directions highlighted” (Hall, N. 2006) was written to bring students and professionals up to date on the

current state-of-the-art in the philosophy of causality. It too makes no mention of “consciousness” or any of its cognates.

It is important not to misunderstand. Neither Dowe nor Hall is in any way at fault. These are fine articles. It is not as if there is some influential publication that they have failed to mention in their literature review. There simply are no sufficiently influential discussions of causality that feature consciousness prominently enough to warrant mention in these survey articles.

Going in the other direction, looking for references to the words “cause,” “causal,” or “causality,” in publications about the nature of consciousness, one does not come up empty handed. Quite the contrary, any publication that promises to give an overview of contemporary theories of consciousness will be compelled to discuss those theories that take consciousness to be a special kind of “access” to certain information processed within a cognitive system. Standard information-theoretic analyses are built upon the assumption that there exist *causal connections* of a content-preserving kind which serve to “carry” information, both from the source to the agent (*via* perception) and between sub-systems within the agent itself (*via* cognition).

Frequent use of the terms, “cause,” and “causal,” notwithstanding, it is rare when consciousness researchers engage in any serious dispute about the *nature* of causality. Most participants to the ontological debates about the nature of consciousness simply take “causality” for granted – even those who vehemently reject a functionalized account of consciousness. All sides seem to concede that the existence and nature of causality is unproblematic. The only disagreement concerns what constitutive role (if any) causality might play in an account of the nature of consciousness.

This paper will challenge the foregoing assumptions that theories of consciousness need not bother with articulating a substantive theory of causality and that theories of causality can safely ignore theories of consciousness. The next four sections of the paper will seek to offer the following: a brief introduction to one long-standing debate about whether the nature of causality is “ontic” or “epistemic” (Section II); an overview of what I will call the “causality revolution” in theories of cognition that began approximately fifty years ago (Section III); an argument to show that information-theoretic accounts of “access consciousness,” which make essential reference to causality, are threatened without a substantive account of what causality *is* and precisely *how* it can secure determinate cognitive content (Section IV); and, finally, an attack on the assumption that it

is possible to give a satisfactory account of the nature of causality in all its dimensions without directly addressing the nature and contribution of phenomenal consciousness (Section V).

2. The Ontological Status of Causality: Epistemic vs. Ontic

There is a long running debate about the nature and ontological status of causality. Consider the classic billiard ball example of cause-and-effect: Momentum is transferred by the direct impact of one rigid object upon another. It is easy to convince oneself that we observe (or at least are justified in inferring) that A is the ontological cause of B, that B was necessitated by A. We will call this kind of causality, “ontic causality” (following Coffa’s 1973 discussion of “ontic” explanation). However reasonable this notion of causality might be, it is controversial. Arguably *the* central conflict regarding the nature of causality not only in the past fifty years but for the past 250 years is whether there is justification for postulating this ontologically robust notion of “ontic causality” or whether we have grounds only to support an ontologically thin but empirically well-supported version of Humean “epistemic causality.”

David Hume (1748/1910) launched an attack not only on the traditional concept of causality, but on the entire metaphysical framework from which that concept arises – including the justifiability of induction. Hume challenged the metaphysical assumptions behind ontic causality by asking what we actually observed in the billiard ball collision case. Hume argued that we do not observe necessity. How could we? All we observe is actuality – our perception of event A followed by our perception of event B. But if we don’t see or hear or touch necessity, then we don’t observe ontic causality either. All we have is our perception of event A followed by our perception of event B. And if we see several instances of a similar kind conjoined, then we observe what Hume calls the “constant conjunction” of A-followed-by-B phenomenon.

If we do not have any empirical evidence of a more metaphysically robust kind of causality, then what is the source of our strong conviction that there is an ontological inevitability that B *must* be followed by A? The source of that conviction, according to Hume, is not a warranted inference from empirically justified premises but a psychological consequence of our propensity to develop habits of mind. It is the product of a sentiment rather than a rational judgment. If we see empirically similar events frequently enough, we form a psychological habit of expectation.

As the debate over causality reached the 20th Century, some of the spirit of Hume's analysis of causality was preserved in the deductive-nomological (or D-N) model of explanation advanced by Carl Hempel and Paul Oppenheim (1948) which carried with it a conception of causality that was considered both empirically responsible enough and ontologically modest enough for the logical empiricist leanings of the day. The D-N or "covering law" account of explanation offered a formal model that captures the way that explanation and prediction constitute parallel practices within the scientific enterprise. Starting with a universal law that captures a regularity in nature in the form of a conditional, the scientist plugs in the initial conditions present at the start of the experiment (the antecedent of the conditional) and generates a prediction by specifying in advance the results of the experiment (the consequent of the conditional). Or, if there is a phenomenon in need of explanation which is an instance of the consequent of the covering-law conditional, and if the conditions specified by the antecedent are confirmed as having been met, then the scientist has an explanation.

The D-N model of explanation, and the clearly "epistemic" conception of causation to which it is committed, dominated discussions about causality for much of the 20th Century. However, as logical empiricism gradually lost its stranglehold in the second half of the 20th Century, a vigorous challenge arose from those who were not shy about seeking a more robust notion of explanation that allows some talk of "ontological necessity" and the resources of genuine causal powers. Wesley Salmon was one of those challenging the received view. Salmon (1990) reviews important literature from the 1970's and '80's and describes the views of those (including himself) who defend an "ontic" conception of explanation.

Along with [Alberto] Coffa, both [Peter] Railton and I endorse the ontic conception. We all maintain that explanations reveal the mechanisms, causal or other, that produce the facts we are trying to explain.

In "Why Ask, 'Why?'" I attempted to respond on behalf of the ontic conception. . . . I suggested that, in addition to purely descriptive knowledge, one would need causal knowledge; recognition of the difference between causal and noncausal laws, the difference between causal processes and pseudo-processes, and the difference between causal interactions and mere spatio-temporal coincidences. At present I would be inclined to phrase the answer somewhat differently in terms of laying bare the underlying mechanisms, but the basic idea seems to me sound. According to the ontic conception, explanatory

knowledge is knowledge of the causal mechanisms, and mechanisms of other types perhaps, that produce the phenomena with which we are concerned. (p. 121)

While the epistemic conception is confined to a descriptive account of what is observable, the ontic conception seeks to ground explanation in the *causal mechanisms* that underlay the observable phenomenon. (Notice, the word ‘causal,’ appears four times in one of Salmon’s sentences explaining ontic explanation.) And yet, even as the ontic conception of explanation was gaining a secure foothold, Salmon, himself, admits that from the start “the main shortcoming of the causal conception . . . was its lack of any adequate analysis of causality.” (Salmon, 1990, p. 119). Salmon took as his challenge the task of giving just such an adequate analysis that was (ontologically) substantial enough to provide genuine explanations for events in the real world, yet rigorous enough to meet the potent challenges advanced by the epistemic view. He describes his view as follows (Salmon, 1990):

The basic question, for this conception, is the status of causality. If we construe causal relations in extrascientific metaphysical terms we will banish explanation from science. If we follow a purely Humean tack, construing causality strictly in terms of constant conjunction (see Mackie 1974), we will make the ontic approach identical with the received view. The approach I adopted in “Why Ask, ‘Why?’?” involved an appeal to causal processes and causal interactions. . . . To my great regret, I found no way of carrying out these explications without the use of counterfactual conditionals . . . I have no analysis of counterfactuals, though I do offer a method for testing the truth of the kinds of counterfactuals that are invoked in this context. I am not terribly dissatisfied with this state of affairs, but doubtless other philosophers will not be as easily satisfied on this score (p. 130).

Even after a decade of trying to work out the details, Salmon here expresses worries about his theory of causality being dependent upon counterfactuals. Later, in (Salmon, 1994), he felt that he could get by without depending upon counterfactuals by borrowing on a new analysis offered in Dowe (1992). But working out the details of a defensible version of the ontic conception of explanation – and of the “ontic causality” upon which it depends – is not trivial. Today, debates continue apace regarding how best to characterize causality so that it is responsible to the empirical rigors of scientific inquiry yet has enough ontological bite to do the work that a substantive theory of causality is often asked to do. It is not the purpose of this paper to give the

reader a comprehensive overview of current defenses of ontic causality so as to pass final judgment on the matter. It is enough to make clear that providing an analysis and defense of a *substantial* theory of causality is no trivial task and that nowhere is a substantive theory of causality asked to do more than it is within the revolution that has placed “causality” at the ontological heart of the very nature of cognition. We turn now to the historical developments that lead to the emergence of theories of cognition and consciousness that are fundamentally dependent upon causality. When it comes to the ontological grounding of these theories, we will find that causality is asked to do most of the heavy lifting – and, with few exceptions, it is asked to do so without the accompaniment of anything like a substantive account of what causality *is* and *how* it can carry the load with which it has been burdened.

3. The “Causality” Revolution Across the Cognitive Domain

There are grounds for asserting that *the* most significant ontological development in all of philosophy in the second half of the 20th Century was the emergence of a new orthodoxy concerning the fundamental nature of what I shall call “the three cognitive domains” – the semantic, the epistemological, and the mental. In an extraordinary tour de force, a philosophical perspective that had dominated for over two millennia was, in a scant twenty years, challenged for supremacy by an entirely new paradigm which reconceived the cognitive realm (language, knowledge, and mind) and built it all upon a single metaphysical foundation: “causality.” Traditional internalist (“Cartesian”) assumptions upon which the three domains had previously been built, were jettisoned by researchers from across every related academic discipline who reduced the cognitive realm to information-carrying physical processes that were grounded in causal relations of the relevant semantic, epistemological and mental type. This view has won adherents in sufficient numbers that it is arguably the new “received view,” at least within cognitive science, and possibly within philosophy generally.

This sea change is frequently referred to as a shift from internalist to externalist theories. Internalist properties are the intrinsic properties of the inner workings of a cognitive system, usually (but not always) including a special emphasis on what is immediately accessible to the first-person perspective of the agent. The relevant externalist properties are information-theoretic properties of the cognitive system in their relationship to the (distal) objects in the external

environment, and the structural-functional properties of the interacting physical parts of the system itself.

Starting in the 1960's, causality-based theories in each of the cognitive domains –mind, language, and knowledge – burst onto the scene, launching a tidal wave that would throw the old, Cartesian internalism onto its heels. This section will offer what can only be a very brief overview of the revolution that took place in these three fields. While this will cover territory familiar to philosophers working in these sub-disciplines, it is surprising that this wholesale shift to a causally-grounded externalism is not better known among consciousness researchers. It is that to which we now turn.

The Causality Revolution in Theories of Mind

In the late 1950s and early 1960's Hilary Putnam was the first to defend a functionalist theory of mind. His original version, *machine state functionalism* (1960), was soon followed by refinements added by a phalanx of supporters from which emerged, *causal role functionalism* which quickly became the dominant theory of mind, a title which it probably still holds today. Traditional internalist theories had located the essence of the mental in the intrinsic properties of the cognitive system, often focusing on what was immediately (epistemically) accessible to the cognitive agent. Functionalism located the essence of the mental in the role played by distinct information-theoretic states within the economy of the cognitive system, thus shifting the central focus to the relational properties of the system. On this theory, the identity conditions for any given mental state are determined by the *causal relationships* that bear among: (1) the information carried by its sensory organs in interaction with its environment, (2) the other mental states with which it interacts, and (3) the overt behavior that it produces.

Consider the belief “There is a beech tree” which consists of both a propositional content and a propositional attitude. At the risk of over-simplifying, we can say that the *content* will be fixed by virtue of the *causal* connections the agent bears to its environment and to other language speakers and the propositional *attitude* (i.e., the fact that it is a *belief* rather than a hope or a fear) will be fixed by virtue of the *causal* relations it bears to other mental states and the *causal* influence it has on the system's overt behavior in the world. All these elements constitute the “causal role” that the mental state plays in the organism's economy. On this theory, to be a mental state just is to be a thing that plays a particular causal role within a cognitive system.

The version of functionalism just described has become the dominant version. Early on, there was disagreement among functionalists regarding exactly what determines the *content* of mental states. It was agreed that a constitutive element of a mental state was the causal role it played in determining the behavior of the organism. There was disagreement, however, regarding the semantic theory that determined the reference of terms in a language (either a language of thought or a natural language). Jerry Fodor (1980), for example, insisted that an internalist semantics was required for a plausible explanation of the cognition that controls human behavior. However, Fodor was very much in the minority, and later, by his (1995) even he conceded that a strong dose of “externalist” semantics (appealing to causal connections bearing between speakers and objects in the environment) was required to flesh out a plausible functionalist theory of mind.

It is actually somewhat surprising, that causality-based theories of mind broke out of the gate before causality-based theories of language were fully developed. It took time to work out the details of a causal semantic theory which could then be fully integrated with the causal role version of functionalism. It is to the revolution in semantics that we now turn.

The Causality Revolution in Semantics

By the mid-to-late 1960’s, the causality revolution was in the air and it was inevitable that a new externalist theory of language would emerge that would complete the picture begun in theories of mind. How is it that words of English come to refer to the objects that they do? Up to this point, a traditional internalist view about how natural languages work was little questioned and was most famously articulated by John Locke in his “way of ideas” (*Essay* 1690). He argued that language speakers had “ideas” in their minds which they associated with a particular term. Those ideas were the descriptive properties that counted as the meaning of the word (its intension). The extension of the term was the set of all objects that possessed the associated properties. For many people this was not just the best theory, it was the only conceivable theory. How else could reference possibly be fixed? The meanings we attach to words are conventions – they are whatever we say they are. And if we all say that what we mean by the term, ‘gold,’ is all and only those things that are “yellow, shiny and malleable” then surely that is what fixes its reference. How could it fail to be so?

Well, it could indeed fail to be so – or so argued Saul Kripke (1971), Hilary Putnam (1970, 1975), and Keith Donnellan (1966). Within a period of just a few years in the late 1960’s and

early 1970's, all three of these philosophers (working independently) produced surprisingly compelling counter-examples to the Lockean definite description theory, and advanced a new theory of reference, often called the "causal theory of reference" (also called the "new theory" and the "direct theory" of reference). On this externalist analysis of reference, it is not the conscious "ideas" floating in the speaker's head that fixes reference – those ideas may be wholly confused. It is rather the *causal connections* between the speaker's use of a term and the objects in the world that causally regulate its use. The semantically relevant causal network will include not only the present speaker (say, *me*), but other members of my linguistic community. Along with Hilary Putnam (who offered this example), I can't tell the difference between a beech and an elm. Yet, when I use the term 'beech' I refer to all and only those objects in any possible world that expert botanists know to be *bona fide* members of the genus "Fagus." What fixes the reference of my term 'beech' is *not* the description floating in my head (which is hopelessly vague), but the actual biological genus that lies at the end of the *causal chains* of the relevant linguistic uses of the term. And while the theory was originally applied only to natural kind terms and proper names, it was soon being used to account for the semantic nature of a broad range of linguistic expressions (Adams, F. et. al. 1992). The monopoly long held by internalist views of language rooted in the conscious states of language speakers was quickly broken by an externalist view of language rooted in ontologically respectable causal connections of a "semantically relevant" kind.

A causality-centered semantic theory is the foundation upon which a causality-based conception of the entire cognitive domain can be built. The semantic theory provides the cognitive content that will become the objects upon which the mind will operate and the information which will be the knowledge-system's defining achievement. It is to causality-based theories of knowledge to which we now turn.

The Causality Revolution in Epistemology

In 1967 Alan Goldman (1967) advanced a simple, intuitive account of knowledge, which he called "the causal theory of knowing." During the next decade, related externalist views were elaborated and expanded by a variety of philosophers. The most influential version that emerged from that early period is the reliable process theory of knowing which Goldman defended in 1975.

Reliabilist theories hold that the feature that distinguishes knowledge (or justification) from mere true belief is *not* some piece of evidence that is internally accessible to the conscious mind of the

agent, but is the reliability of the organism’s “belief-forming mechanisms.” If the process that produced the true belief is reliable – a fact that may be wholly inaccessible to the agent herself – then the belief counts as an instance of knowledge (or of justified belief, if it is a reliabilist theory of justification).

This theory had immediate and wide-reaching appeal within the (then rather insular) philosophical community of epistemologists because it offered an account of knowledge which handled a range of cases that had plagued traditional theories. Its reception was even more enthusiastic and more pervasive within those disciplines that contribute to the cognitive sciences – especially by the time Fred Dretske (1970, 1981) had developed a sophisticated information-theoretic externalist theory built upon the seminal work of Claude Shannon (1948) on the mathematical theory of information. Dretske grounded knowledge in the information-carrying signals with which the universe is replete. Information processing systems – whether humans, worms, and possibly even thermostats – function as receivers when they are able to extract information from those signals. While the theory is very complex by the time all of the details (and qualifications) are enumerated, the crux of the matter is that knowledge consists in the proper reception of information being carried by epistemically relevant *causal processes*. “Causality” again lies at the center of the picture.

This basic externalist picture of cognition – including semantics, mind, and epistemology – was largely in place by the early 1970’s and would seem to complete the causality-based account of cognition. However, there remained a gap in the account of “the mental,” because at that time in history no one had attempted a fully articulated account of consciousness. The 1980’s and ‘90s brought an impressive onslaught of causality-based theories of consciousness that sought to complete the causality-revolution.

The Causality Revolution in Theories of Consciousness

As we saw previously, functionalism is a theory of mind that ignores the intrinsic, internal properties of mental states (including any first person phenomenal states) and locates cognition solely in relational (information-bearing) external properties. Advocates of this view, when turning to the essential nature of *consciousness*, continued to flesh-out a global, information-theoretic perspective and focused, reasonably enough, on the *causal-role* that conscious states play within the economy of cognitive systems. Bernard Barrs (1988, 1997), for example, offered a

global workspace theory which conceives of consciousness on the rough analogy of a spotlight on a theatrical stage that makes cognitive content immediately accessible to multiple modules within the cognitive system.

Ned Block (1995) drew the now familiar distinction between two kinds of consciousness. “Access consciousness” is a functionalized notion of consciousness for which Baars has offered a paradigmatic example. Block contrasts this with “phenomenal consciousness,” which is the qualitative character of first-person, subjective experiences. A proponent of the causality revolution in cognition has several options when it comes to consciousness. One option is to deny the distinction altogether, insisting that access consciousness is the only kind that exists (Dennett, 1991). A second option is to concede that phenomenal consciousness *exists* (and thus has a place in one’s ontology) but deny that it is in any way constitutive of *cognition*. This is the route taken by epiphenomenalists who, like Jaegwon Kim in his (2005), think that intentionality – and thus cognitive content – can be functionalized and thereby subsumed under the causality revolution, but that phenomenal consciousness is a loose wheel, disengaged from cognition proper. A supposed third option, representationism (Tye, 1994), advertizes itself as a causal reduction of phenomenal consciousness itself. This is a semantic move offering a new definition of phenomenal consciousness and is not, for the sake of this paper, ontologically distinct from the first option described above.

For those who embrace the causality revolution, the most common account of consciousness is to describe the causal role that consciousness plays within the economy of the cognitive system (thus fitting nicely into the overall causal reduction). This leaves no explanatory work for phenomenal consciousness to do, so it is either left out of the picture altogether (viz., eliminativism) or is reduced to an “epiphenomenal” irrelevancy. One of the most popular ways of describing it – in the spirit of the “causality revolution” – is to say that mental states will be “conscious” in the cognitively relevant sense by virtue of the content-carrying causal connections they bear to other parts of the cognitive system, thus allowing them to “broadcast” their content to those sub-systems.

This completes a quick survey of the way that causality lays at the foundation of an externalist view of cognition. While there are many challenges that a causal reduction of cognition must overcome, we will consider just two: one involving cognitive content (which threatens externalist theories of both language and the mind), and the other involving knowledge.

4. Challenges to Causality-Dependent Cognition

As we have seen from the “causality trifecta,” each of the three cognitive categories has made *causality* the tent pole around which the theory is constructed. While these theories continue to enjoy broad support, each theory has been attacked for what can best be described as “the failure to demonstrate that causality has the resources to do the job (semantic, mental, and epistemological) that its theory requires of it.” The debates involved here are, admittedly, complex and only a brief overview is possible. However, I hope to show that important objections have been raised that rest on the impotence of causality to do the work that the theory claims for it. The goal of this survey is to substantiate the claim that causality-dependent theories of consciousness have not paid full fare until they have defended a substantive theory of causality and shown how it is capable of supporting the entire cognitive realm.

Causality-Dependent Epistemology and the Generality Problem

The intuitive idea behind most externalist epistemological theories, including reliabilism, is that one’s belief that P (if it is to be knowledge) should be causally hooked up to *the fact that P* in the right sort of way. There should be a *causal connection* between the cognitive mechanisms that produce the belief, and the object of that belief. Reliable cognitive systems will be causally hooked up to their environment in such a way that changes in the environment will be causally transmitted to the cognitive systems and will be reflected in changes in the beliefs generated. All the theory requires, then, is a plausible way to identify the relevant “reliable processes of belief-formation.”

This proves to be something less than perfectly straightforward. How are “processes” to be individuated? Consider the case where you look at a maple tree and come to believe P: “There is a maple tree.” What is the relevant “process” whose reliability we must measure? Your coming to believe P is an individual *event*. It is a *token* processing event, not a type. The theory must identify the relevant *type*. We could define the type so narrowly that it only has one instance, the instance of your coming to believe P just now. We could define the process-type that narrowly: The “maple-tree-identification-that-happens-under-precisely-the-conditions-obtaining-on-January 15, 2011 at 11:03am” belief-forming-type. If we individuate the process that narrowly, there will be only one belief ever generated by *that* belief-forming-mechanism and of course it will be 100%

reliable because the only time that process was ever instantiated it resulted in a true belief. If we define processes that narrowly, there will be as many types as there are belief-tokens. Worse, every true-belief will be justified and every false belief will be unjustified – the very distinction between knowledge and mere true belief will collapse. So we must define the relevant belief-forming-process more broadly. But what *is* the relevant process? The “visual identification of an object made in daylight” process? Or the “visual identification of flora made in daylight” process? Or the “visual identification of a tree made in daylight” process? Or the “visual identification of a deciduous tree made in any light” process?

This difficulty is the “generality problem” which was first advanced by Richard Feldman (1985) and further elaborated by Feldman and Conee (1998). The problem is that the individual event in which you judged that you saw a maple tree on January 15, is an instance of an *infinite number* of different cognitive-processing-types. It is no small task to settle which process is the epistemically relevant one. Let’s assume that we can come to agree on the “relevant-belief-forming-process-type” – call it T. By virtue of *what* exactly will it have qualified as being the correct one? As an externalist reduction of epistemology, it was advertised that “causal-connections” of the relevant type were going to settle the matter, replacing the rejected “internalist” conceptions of justification. But there are too many causal connections. Each of the infinite number of reliable processing types is causally connected to our belief that P. We want only the epistemically relevant process-types. At this point, though, the only obvious grounds for choosing one process-type over another is by appeal to our internalist epistemological intuitions – which is simply a sneaky way of dressing up the old internalist theory in new externalist clothes. Causality is supposed to be doing the epistemological heavy lifting here (it is a reductive theory after all), but we don’t seem to have a conception of causal connections that can do the work – regardless of how ontically robust that conception is.

Causality-Dependent Content: The Indeterminacy Problem for Semantic & Mental Content

The second problem we will consider is a threat to the very notion of cognitive content –the fundamental concept at the center of both externalist semantic theories and externalist theories of mind. In the previous section, we saw that it wasn’t obvious that causality possessed the resources to determinately fix the process-type to which a reliabilist epistemological theory would be calibrated. Without such a fix, it will be indeterminate what the theory actually is. In this section

we will see how causality seems also to lack the resources to determinately fix cognitive content – the very commodity in which the information-theoretic systems are supposed to trade. Without such a fix, there will be no fact of the matter regarding what information is being communicated (received, sent, or processed), and so it will be indeterminate what proposition is expressed by any sentence (semantics) or contemplated by any thought (mind).

W.V.O. Quine (1960) raised the problem of content indeterminacy in both his doctrines of the “indeterminacy of translation” and the “inscrutability of reference.” He offers the famous case of the field linguist coming across a tribe who point at a rabbit and say, “gavagai.” The only data available to fix the reference of the term is the behavior of the speaker, but that behavior is compatible with interpreting ‘gavagai,’ as meaning any number of different things other than what *we* consider to be a plain old rabbit – including “undetached rabbit-part” and “rabbit time-slice.” Quine’s perspective on both language and the mind was more behaviorist than the functionalist theories we have been considering, but he was one of the leading lights in the “naturalization” of the entire cognitive domain and was generally in favor of physicalist reductions whenever they were compatible with his equally strong commitment to radical empiricism. Quine’s view is too nuanced to describe here, but it is important to recognize that in spite of Quine’s firm commitment to reductive physicalism, he conceded that naturalizing semantics in this way would require that we leave *meaning*, as traditionally conceived, out of our ontology. Since our terms don’t have fixed, determinate content, then our theories better do without them. Quine was able to embrace this kind of semantic eliminativism, but most advocates of the causality revolution are not so willing. In an information-theoretic system, so much theorizing appeals to the notion of “content,” and presupposes that it refers to something determinate, that living without that assumption is too bitter a pill to swallow for most defenders of the theory.

Daniel Dennett, whose own theory is more functionalist than behaviorist, concurs with the indeterminacy assessment.

But I also maintain that when these objective patterns fall short of perfection, as they always must, there will be uninterpretable gaps; it is always possible in principle for rival intentional stance interpretations of those patterns to tie for first place, so that no further fact could settle what the intentional system in question *really* believed. (Dennett, 1987, p. 40)

Dennett's position on the ontological status of beliefs (and other mental states) is instrumentalist. To be an instrumentalist is to agree with the eliminativist about the ontological question (i.e., meanings, beliefs, and contents don't *really* exist) but to hold that in spite of that there is reason to continue to use the terms ('meaning,' 'belief,' and 'content') because they have instrumental value. Semantics is as it may be, but at the ontological level, Dennett joins Quine in admitting that causality does indeed lack the resources to produce determinate cognitive content.

Orthodox defenders of the causality revolution must resist the hardcore Quine-Dennett conclusion since it is radically revisionist and thus fails to be a genuine "reduction" as we understand the term. Knowing this, enemies of the causality revolution use the threat of indeterminacy that arises from these theories as grounds for claiming that *phenomenal consciousness* not only exists, but is essential to the production of determinate cognitive content. This position is one popular plank of a theory called, "phenomenal intentionality," (Loar 1987) which had been explicitly defended by very few until the 1990s (although many would say implicitly assumed by Husserl and Brentano long before that). In the past two decades, however, interest in the view has mushroomed and we might now be justified in calling it a "movement" – if still a modest one.¹ Terry Horgan and Uriah Kriegel, two of the champions of this view, discuss externalist theories that make use of "tracking relations." Tracking relations will be instantiated in causal chains that enable a cognitive system to be sensitive to the presence or absence of certain features of their environment. Even if you add evolutionary forces that are designed to fine-tune various tracking systems, the resources are still inadequate for fixing content. They say:

. . . nothing other than phenomenal character can secure content determinacy. In particular, tracking relations cannot account for this determinacy – not even when teleologically augmented. Whenever an internal state bears tracking relations to rabbits or Phosphorus, it also bears them to collections of undetached rabbit parts or Hesperus. And as Fodor (1984, 1990) argued long ago, whenever tracking a property F is adaptive for an organism, it is also and equally adaptive for it to track any property coextensive with F. Evolutionary benefit cannot discriminate between coextensive properties (let alone necessarily coextensive ones).

¹ For my own discussion of what is right about phenomenal intentionality, see my (Forthcoming) "Machine Intentionality, the Moral Status of Machines, and the Composition Problem" where I defend an original theory of intentionality ("moral status intentionality"). There I admit that the phenomenally conscious states of cognitive agents may well be the only grounds for the attribution of "moral status," in which case moral status intentionality and phenomenal intentionality will agree about the conditions necessary for genuine intentional content.

The point made here by Horgan and Kriegel is an important one. Probably the most popular strategy employed by defenders of the causality revolution who recognize that they need to solve the indeterminacy problems of both content and knowledge, is to make a teleological appeal to natural selection (Millikan, 1989). The objection by Horgan and Kriegel claims that the prospects of reducing “teleology” to a causal foundation seems no more promising than the prospects in any of the other areas of cognition.

Summary Comments on “Causality-Dependent Cognition & Consciousness”

Thus far, we have explored the role that causality plays in reductive accounts of cognition – from semantics and epistemology to theories of mind and consciousness – and offered reasons for thinking that an adequate case as yet to be made for this view. Defenders of this view rarely, if ever, attempt to give a substantive account of the nature causality or a satisfying explanation of how their concept of causality (whatever it might be) can do the work required of it. This is a call for defenders of the causality revolution to take up the challenge and explain precisely what causality is and how it can accomplish all that an externalist reduction of cognition requires of it.

We now leave behind “causality-dependent” theories of cognition and consciousness and consider the possibility that the tables might be turned and that we might discover a “consciousness-dependent” theory of causality.

5. Consciousness-Dependent Causality

We now come to a fundamental turning point. Thus far, we have been considering the family of theories according to which cognition and consciousness are ultimately grounded in causality. Our analysis has shown that this hypothesis is threatened by a vicious circularity that relies either upon irreducibly internalist elements in the very notion of causality (if *epistemic* causality is assumed) and/or internalist judgments in cashing-out the concept of causal connections “of the appropriate type” if *ontic or epistemic* causality is appealed to.

We now turn to consider a hypothesis that reverses the order of ontological priority between causality and consciousness: “Causality (or rather, one type of causality) reduces to conscious, observer-dependent interactions of the appropriate type.” I will argue that there exists a *consciousness-dependent causality* – a species of epistemic causality – that is not given its due by either side in the traditional causality debate. One obstacle that this conception of causality must

overcome is the unexamined assumption that there is symmetry between the ontic and the epistemic theories of causality. On this assumption, there is only one kind of causality and whichever side wins the day, the other side is altogether vanquished.

Against this assumption, I contend that the universe contains two kinds of causality. While I have argued above that there is a substantive debate about whether ontic causality actually exists, I have, to this point, resisted the temptation to take sides on the question. I am now prepared to admit that I do include ontic causality among my own ontological commitments. However, just because ontic causality does exist, it does not follow that epistemic causality does not exist. The first step to recognizing that epistemic causality deserves a place in our ontology (even if ontic causality also deserves such a place) is to acknowledge that there is an asymmetry in the traditional debate between ontic and epistemic causality.

The traditional assumption is that there is symmetry in the competition between ontic and epistemic causality. Each offers an alternative theory of one and the same phenomena. Consider, for example, the classic billiard ball scenario. We judge that there is a relationship that bears between the first billiard ball sitting on the table, and the second ball that strikes it. The question is: Which property most accurately captures the nature of that relationship? This way of posing the question obscures the asymmetry. While the defenders of the epistemic theory do indeed insist that their theory does and the ontic theory does not describe a relationship that exists between the two billiard balls, the defenders of the ontic theory do *not* insist that the relationship picked out by the epistemic theory fails to hold between the two events. *No one* denies – no one could deny – that there exists a “constant conjunction” of perceived events of type-A (“ball 1 striking ball 2”-type events) followed by perceived events of type-B (“ball 2 moving at a predictable angle & with predictable velocity”-type events) in normal billiard ball situations. This is an entirely innocuous claim, ontologically weak enough to be uncontroversial. Thus, defenders of ontic causality do not deny the *existence* of epistemic causality, they simply deny that it is *real* causality, properly so called. Neither side can deny that the relationship picked out by epistemic causality exists. The only genuine ontological dispute is whether there exists a second, ontologically more robust relationship, of *ontic* causality. Virtually everyone who believes in the existence of ontic causality, also believes that it alone is the proper referent of the term, ‘causality.’ Traditionally, the only defenders of epistemic causality are those who do not find plausible support for the existence of

ontic causality, and so are left with the weaker notion of epistemic causality as the only possible referent of the term.

I am going to go against custom and defend the existence of *both* ontic and epistemic causality. It is undeniable that we perceive Humean constant conjunctions of type-A events followed by type-B events. What is in contention is *not* their existence, but the appropriateness of attributing the property of “causality” to them. It is not a question of whether to include these things in our ontology – we all do that anyway – it is only a matter of whether we are justified in including them *under the description*, ‘causality.’ It is arguments in defense of *that* conclusion that I am now prepared to offer.

Phenomenal causality is REAL!

Against all the traditional views described above and against the entire weight of philosophical tradition, I want to suggest that it is one of the central cognitive achievements of the human race that we not only *can*, but at times actually *do* – as a matter of contingent fact – privilege the epistemic over the ontic, even when we are fully convinced that the ontic is real. What I will refer to as the “priority of the epistemic” is not the priority born of an antirealist rejection of or a Humean skepticism about the existence of the ontic domain. Rather, it is the priority conferred by the judgment of a conscious agent who makes her immediate phenomenal experiences the subject of her cognitive interest rather than any ontic reality. It is the priority bestowed by the conscious judgment we make when we determine that the *objects* of our discourse, the *truth-conditions* of our utterances, and the *causal relations* that obtain between the objects that populate a particular environment, will be determined *not* by any ontic reality but by an epistemic reality alone. We call this an “epistemic” reality because it is the phenomenal world *as it is perceived* by the epistemic (i.e., cognitive and sensory) faculties of a conscious agent rather than any ontic world that exists independent of the activities of conscious agents.

This account is admittedly abstract, and cries out for a concrete example to show precisely how the rubber meets the road. Only a brief example is possible, here, but the nature of the thesis should become apparent. Consider the case where you and I are playing a video game. We both know that we are sitting in a room that includes a table, a gaming computer and controllers, two chairs and a lamp. There are no tennis balls, no tennis racquets, and certainly no tennis court in the room. Yet, we are both looking at the television screen and speaking about the objects on the

screen. I say: “Hit the ball already.” At the ontic level, there is only a two dimensional screen consisting of liquid crystal pixels which alter their color every fraction of a second. The two of us viewing the flat screen do not “see” a two-dimensional plane of flashing lights. Instead, we see a three-dimensional world, which includes a tennis court into which three-dimensional tennis balls can be served.

You and I, ignoring the so-called “real” mind-independent objects in the “real” mind-independent room, speak only of the reality that exists on the screen. But of course, the reality that exists *literally* on the screen is nothing but flashing colored lights in a two-dimensional array. The only place there exists any three-dimensional tennis court is, as we might say, in our minds. More formally, we might say that there is a court, balls, and racquets “from the epistemic perspective of” the participants.

It is important to pause a moment and note that this ability of the human visual system to take two-dimensional input from the monitor of a video game and convert that into the experience of a three-dimensional world is hardly unique to video games. It is precisely the same ability our visual system exercises every time we experience the three-dimensionality of the physical world. Vision itself results when the brain takes the tiny amount of data projected – upside down – onto the two-dimensional array of the eye’s retina and generates from that modest input a bounty of perceptual output in the form of the rich three-dimensional world that we experience. The *inverse problem* is the name of the challenge that the visual system must overcome in reconstructing a 3D world out of such meager 2D data; visual illusions are one of the by-products when the system fails. (See Anderson 2005 for an undergraduate-level introduction to these basic issues in vision science.)

Let us return to the virtual tennis match and our discourse about it. You say: “I just hit a great serve out wide to your forehand.” Now if that sentence is interpreted *ontically* as a claim about physical, mind-independent objects, then the sentence is patently false. There are no tennis rackets in the room and no one hit a serve. The sentence is equally false if we interpret it as an ontic claim about the television screen. Again, there are only flashing liquid crystals, there are no tennis balls. Yet, it is a contentious, philosophically motivated stand to take to insist that every sentence either of us utters is false for the entire two hours we are playing the game, simply because there is no *physical* tennis ball to make our sentences true. Surely the only plausible semantic position is to interpret the sentence to be a claim about the *virtual* tennis ball and the

virtual court – which objects only exist from the epistemic perspective of the human beings watching the screen.

If this interpretation is correct, as I have argued in Anderson 2009, then the sentence you uttered is true, and what *made it true* is the fact that there exists a virtual tennis ball – epistemically accessible to the two of us – that did indeed go out “wide to my forehand.” It is real enough to make your sentence true but its reality is grounded at the epistemic level not the ontic level; its reality is literally constituted by the operations of the cognitive and sensory faculties of the two of us playing the game. In a world that contains only a screen and flashing colored lights, but no human beings, there would exist no tennis ball – even if the screen in question flashed tennis-ball-like images just as in the actual world. It is not the flashing lights alone that constitute the virtual tennis ball, it is the operations of your and my cognitive and sensory faculties and the shared, even public, “empirical world” that they generate.

We return, now, to the implications of all of this for the nature of causality. When I say I hit a serve out wide, I could easily have used the language of causality and said “The elegant and powerful way that you swung your racket *caused* the tennis ball to go 124 mph out wide to my forehand.” Since there are no reasonable grounds for saying that the sentence is false, we must ask what conception of causality is being invoked by virtue of which this is a true causal claim? Is it ontic causality or epistemic causality that is being invoked? I claim it is epistemic causality. And, because in this case it is grounded in the phenomenally conscious states of the relevant agents, I will call this particular kind of epistemic causality, “phenomenal causality.”

Admittedly, there are defenders of global externalism who will want to take the position that sentences about virtual tennis balls can be true, but will, nonetheless, want to dismiss talk of “phenomenal causality” and confine the causal relationships to those physical relationships that bear between the “real” objects in the room, including the swinging of the controller, the response in the computer software, the state of the “tennis-ball-subroutine” in the computer program, the coordinated flashing LCD’s on the monitor, and the like. Instead of defining the tennis racquet in experiential terms, so that it is the product of our cognitive faculties, they will instead say that the term, ‘tennis ball,’ refers to some combination of the physical systems just enumerated.

It is impossible to offer, here, a complete rebuttal to this externalist objection to the virtual tennis case (although a more detailed rebuttal is found in Anderson 2009). Just one brief argument will suffice for this discussion. Imagine a possible world that has all of the physical systems

described above – computer, controllers, monitor, etc. – behaving exactly as they are behaving while my friend and I play “virtual tennis” – yet lacking only the presence of human beings. This world we are imagining does not and never has included any human beings with cognitive faculties like our own. Since there is no human to be swinging the controller which generates changes in colors of the flashing LCDs, we can imagine instead a machine of some kind moving the two controllers in exactly the same motions that my friend and I move our controllers. How is this happening? It doesn’t matter. Imagine that the arrangement of molecules in the computer and other machines involved is simply a bizarre quantum accident. Yet, all of the physical, causal relations continue to obtain. An externalist, who seeks to identify the virtual tennis balls with something physical (e.g., flashing lights or sub-routines in the computer program) will, in the case just imagined, have to say that the sentence “A serve was hit out wide to the right at 124 mph” will come out true. But in the world we’ve imagined, a world without creatures like ourselves who *experience* 3D tennis balls and racquets, there is no justification whatsoever, to describe the flashing lights on the monitor or the on-off switches in the computer program as representing a 3D tennis ball. That interpretation requires that we smuggle in the very idea of creatures with cognitive and sensory faculties like our own. It requires that we presuppose an “epistemic perspective” – the very thing that externalism claims to replace with “causal connections of the appropriate type.” There is no escaping the “priority of the epistemic.” We can’t help but interpret the physical goings on in terms of the phenomenal experiences we *would have* were we to be in that world, viewing the screen.

The only reasonable position to take on the semantics of virtual tennis is that the truth-conditions of such utterances are determined *not* by any ontic causal connections bearing between physical objects in the universe, but rather by epistemic causal connections bearing between phenomenal objects that exist only within the epistemic perspective shared by conscious agents. Phenomenal causality is *real*, but it can only exist in universes that contain creatures capable of phenomenal consciousness.

The place of phenomenal causality within a broader theory of cognition

In the few pages that remain, it is impossible to give an adequate explanation and defense of phenomenal causality. I will, however, address possible confusions that might arise concerning

the claims just made for it, consider objections that might be raised, and, most importantly, locate phenomenal causality within a broader internalist theory of cognition.

First, remember that *epistemic* causality is the Humean concept of causality (explained in Section II) that is the sister-concept to what Salmon called “the Epistemic Conception” of explanation, which he considered the received view and against which he posited “the Ontic Conception” of explanation and its sister-concept of *ontic* causality. Hume argues that the only metaphysical system that is justifiable is an exclusively epistemic one. He denies that we are justified in believing in a mind-independent reality that (ontically) *causes* us to have first-person experiences; our only justification (such as it is), comes from the irresistible cognitive habit we have of assuming that our future subjective experiences will be like our past experiences. Inferences, then, are from the epistemic (how things “seem” from the 1st person cognitive perspective of the agent), to the epistemic (further such “seemings”); they are *not* from the epistemic to the ontic (as epistemological foundationalists would have it).

It is this historical notion of epistemic causality that I am recruiting as a resource to help fill-out a comprehensive internalist theory of cognition that is an alternative to global externalism. *Phenomenal causality* is a species of empirical causality that is specifically grounded in the phenomenal properties of cognitive agents. (This leaves open the possibility for other kinds of empirical causality that are also grounded in the epistemic faculties of cognitive agents but which are not specifically tied to phenomenal consciousness.) Phenomenal causality, is not a lone metaphysical posit – if it were, it would be a strange ontological dangler. Instead, it is one element within a theory of cognition which invokes two distinct types of cognitive content. Focusing on semantics (one could as easily focus on the content of mental states), phenomenal causality is one element in a semantic theory that claims there is a language that we all speak whose truth-conditions are grounded not in the states of mind-independent (“external”) objects but rather in the (equally real) phenomenally conscious states of cognitive agents. If I were to claim that this “epistemic” language was the only language we were capable of speaking, then I would be a global antirealist (like Kant, Putnam 1981, Dummett, etc.) who denies that we can think or speak about a mind-independent, ontic world. I am no antirealist. I am a card-carrying metaphysical realist (and a believer in ontic causality). The signature claim I am making here is a semantic one, not a metaphysical one. I claim that we speak two languages, not just one (i.e., two versions of English, not just one) – a view I call *semantic dualism*. Sometimes when I utter the sentence “There is a tree

in the quad” I utter it with realist force investing it with *realist truth-conditions* which means that it makes claims about mind-independent reality; most utterances of that sentence, however, will have *epistemic* (typically called “anti-realist”) *truth-conditions*, which means that it makes claims about how things *seem* (ideally and in the long run) from the first-person, phenomenally conscious perspective of cognitive agents. There are many arguments justifying this semantic theory (Anderson 1995, 2002), but the basic intuition is quite simple. In some contexts (especially philosophical ones), I intend to make a robust metaphysical claim: For example when I pound on the tree and insist “This is not just a phenomenological tree, as Berkeley believed, it is a mind-independent tree.” In that speech context I utter the sentence with realist truth-conditions thereby making a claim about a mind-independent tree. However, in normal contexts, I do not intend to be asserting that Berkeley was wrong (in believing only in a mental world) nor that he was right. In normal contexts, I don’t intend to be making *any* deep metaphysical claims about the fundamental *nature* of the tree, I only intend to speak of the empirical tree that can be seen and touched and that is otherwise accessible to the epistemic perspective of cognitive agents. We all agree about the properties of the empirical tree; it is *that* tree (and only that tree) about which I presently want to speak.

If *semantic dualism* is true, then there are two worlds of which we think and speak: a physical world and a phenomenal world. While it is not possible to repeat the arguments previously advanced in defense of this claim (Anderson 1995, 2002), one should be able to see that if the semantic theory is true, then there would be *prima facie* grounds for interpreting the term, ‘causality,’ as *ontic causality* for those sentences requiring a realist interpretation and as *epistemic (or phenomenal) causality* for those sentences interpreted as antirealist claims about how things seem from the epistemic perspective of cognitive agents.

We can now return to the virtual tennis example and locate that in its broader context. Where one might have thought that it was only in discussions of video games that phenomenal causality would have a place, it turns out (if semantic dualism is true) that *many* sentences we utter have epistemic (not ontic) truth-conditions and thus engage an epistemic not an ontic notion of causality.

Consider one final objection to the virtual tennis example. For those not persuaded by semantic dualism, the virtual tennis example remains the central argument in defense of

phenomenal causality. It has been suggested² that the virtual tennis argument can be defeated by embracing a fictionalist interpretation of virtual tennis discourse, thereby accounting for our “talk” of tennis balls and tennis racquets without committing ourselves to the literal *truth* of sentences about virtual racquets and virtual balls. “Fictionalism” is a newly popular name for a family of theories (many of whose members have been around for a long time) which offer various ways of interpreting discourse that requires some kind of non-literal interpretation. There are a half-dozen or more different versions of fictionalism, but for our present discussion it is important to distinguish two broad types. One type borrows from “pretense theory” in debates about the nature of fiction writing. In a fictionalized novel about my life the author might *say* that I hit an ace to my opponent’s backhand to win a tournament, even though in real life I never did. This is like a game of *pretend*. It is merely a story. “Let’s pretend something is true that we know to be false.” A second type of fictionalism is more akin to operationalism with respect to scientific theories. On this account, the proper interpretation of the meaning of ‘electron’-statements is *not* to make a claim about the existence of tiny theoretical entities called, “electrons,” it is rather to make a claim about something else – in this case, a claim about what traces will be observable in a cloud chamber. This is *not* to pretend that electrons exist when they really don’t; it is rather to make a claim about what is real and true about the operations of the cloud chamber. It is no less a claim about objective reality – it is just a different (less controversial) claim.

Discourse about virtual environments is more like operationalism than like pretense-theory. When you say “I hit a hard serve to your backhand to win the tournament” – you are *not* saying that we should pretend you did something that you didn’t really do; you are saying that you *really did do* something that anyone with the proper sensible and cognitive faculties could have verified – if only they were watching the computer monitor. In other words, statements about virtual tennis matches make genuine truth-claims (presupposing the most robust kind of *correspondence truth*). They are simply making truth claims about what is observable from the epistemic perspective of human beings rather than about mind-independent objects.

6. Conclusions

It has not been the goal of this paper to vanquish causal theories of cognition and consciousness. That would take a good deal more than what I have offered here. It has been my goal, however, to

² Thanks to an anonymous reviewer for this objection.

try to slow down the “causality juggernaut” and to raise serious questions about what I believe to be a naïve reliance upon “causality” to resolve puzzles about the nature of cognition and consciousness that are not so easily solved. I have also sought to champion the “priority of the epistemic.” Our thoughts, attitudes, and utterances are *not* always about “the ultimate nature of mind-independent reality.” They are often about the “perceived world” generated by our sensible and cognitive faculties – regardless of how things stand with respect to mind-independent reality. Thoughts and utterances about *that* world presuppose an epistemic notion of causality. We need a theory of cognition and consciousness that takes both the ontic and the epistemic dimensions of reality seriously.

REFERENCES

- Adams, F., Stecker, R. and Fuller, G. (1992) The Semantics of Thought, *Pacific Philosophical Quarterly*, **73**, pp. 375-389.
- Anderson, D. L. (1995) A Dogma of Metaphysical Realism, *American Philosophical Quarterly*, pp. **32**, 1-11.
- Anderson, D.L. (2002) Why God is not a Semantic Realist, in Alston W. P. (ed.) *Realism and Antirealism*, Ithica, NY: Cornell University Press.
- Anderson, D.L. (2005) The Science of Vision: An Introduction, *The Mind Project* [Online] <http://www.mind.ilstu.edu/curriculum/modOverview.php?modGUI=204> [24 July 2011]
- Anderson, D.L. (2008) Consciousness and Realism, *Journal of Consciousness Studies*, **14** (1), pp. 1-17.
- Anderson, D.L. (2009) A Semantics for Virtual Environments and the Ontological Status of Virtual Objects, *American Philosophical Association Newsletter on Philosophy & Computers*, **9** (1), pp. 15-19.
- Anderson, D. L. (Forthcoming) Machine Intentionality, the Moral Status of Machines, and the Composition Problem, in V. C. Müller (ed.) *The Philosophy and Theory of Artificial Intelligence*, New York: Springer.
- Baars, B. (1988) *A Cognitive Theory of Consciousness*, Cambridge: Cambridge University Press.
- Baars, B. (1997) *In the Theatre of Consciousness*, Oxford: Oxford University Press.

- Blackburn, T. (1988) The Elusiveness of Reference, *Midwest Studies in Philosophy Vol XII: Realism and Antirealism*, Minneapolis: University of Minnesota Press, pp. 179-194.
- Block, N. (1995) On a confusion about the function of consciousness, *Behavioral and Brain Sciences*, **18**, pp. 227-47.
- Boghossian, P. (1989) Content and Self-Knowledge, *Philosophical Topics*, **17**, pp. 5-26.
- Coffa, J. A. (1973) *The Foundations of Inductive Explanation*, Doctoral Dissertation, University of Pittsburgh.
- Dennett, D. (1987) *The Intentional Stance*, Cambridge, MA: The MIT Press.
- Dennett, D. (1991) *Consciousness Explained*, Boston: Little, Brown and Company.
- Donnellan, K. S. (1966) Reference and Definite Descriptions, *The Philosophical Review*, **75**, pp. 281–304.
- Dowe, J. (1992) Process Causality and Asymmetry, *Erkenntnis*, **37**, pp. 179-96.
- Dowe, J. (2009) Causal Processess, *Stanford Encyclopedia of Philosophy*. [Online], <http://plato.stanford.edu/entries/causation-process/> [24 July 2011]
- Dretske, F. (1971) Conclusive Reasons, *Australasian Journal of Philosophy*, **49**, pp. 1–22.
- Dretske, F. (1981) *Knowledge and the Flow of Information*, Cambridge, MA: MIT Press.
- Feldman, R. (1985) Reliability and Justification, *Monist*, **68**, pp. 159–174.
- Fodor, J. (1980) Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology, *Behavioral and Brain Sciences* **3**, 63-109.
- Fodor, J. (1995) *The Elm and the Expert: Mentalese and its Semantics*, Cambridge: MIT Press.
- Gettier, E. (1963) Is Justified True Belief Knowledge?, *Analysis* **23**, pp. 121-123,
- Goldman, A. (1967) A Causal Theory of Knowing, *Journal of Philosophy*, **64**, pp. 357-372.
- Goldman, A. (1975) Innate Knowledge, in S.P. Stich, S.P. (ed.) *Innate Ideas*, Berkeley, CA: University of California Press.
- Hempel, C. G., and Oppenheim, P. (1948) Studies in the Logic of Explanation, *Philosophy of Science*, **15**, pp. 135-75.

- Horgan, T., and Tienson, J. (2002) The Intentionality of Phenomenology and the Phenomenology of Intentionality, in Chalmers, D. (ed.) *Philosophy of Mind: Classical and Contemporary Readings*, Oxford: Oxford University Press, pp. 520–533.
- Horgan, T., Graham, G., & Tienson, T., (2007) Consciousness and Intentionality, in Schneider, S. and M. Velmans, M. (eds.) *The Blackwell Companion to Consciousness*, Oxford: Blackwell Publishing LTD, pp. 468-484.
- Hume, D. (1748/1910) *Enquiry Concerning Human Understanding*, Harvard Classics, Volume 37, P.F. Collier & Son, Section VII, Part 2.
- Kim, J. (2005) *Physicalism, Or Something Near Enough*, Princeton, N.J.: Princeton University Press.
- Kripke, S. (1971) Naming and Necessity, in *Semantics of Natural Languages*, Davidson, D. and Harman, G. (eds.), New York: Springer. 1971. Republished as a separate volume, (1980) *Naming and Necessity*. Cambridge: Harvard University Press.
- Locke, J. (1690/1975) *An Essay Concerning Human Understanding*. Nidditch, P. (ed.) Oxford: Oxford Clarendon Press.
- Loar, B. (1987) Subjective Intentionality, *Philosophical Topics*, **15**, pp. 89-124.
- Millikan, R., (1989) Biosemantics, *Journal of Philosophy*, **86**, pp. 281–97.
- Putnam, H. (1960) Minds and Machines, in Hook, S. (ed.) *Dimensions of Mind*, New York: New York University Press.
- Putnam, H. (1970) Is Semantics Possible?, *Metaphilosophy*, **1**, pp. 187-201.
- Putnam, H. (1975) The Meaning of ‘Meaning’, in Gunderson, K. (ed.), *Language, Mind and Knowledge. Minnesota Studies in the Philosophy of Science*, Minneapolis: University of Minnesota Press, pp. 131-193. Reprinted in (1975) *Mind, Language and Reality*, Cambridge: Cambridge University Press, pp. 215-271.
- Putnam, H. (1992) Why Functionalism Didn’t Work, in Earman, J. (ed.) *Inference, Explanation, and Other Frustrations: Essays in the Philosophy of Science*. Berkeley: University of California Press, pp. 255-270.
- Putnam, H. (1981) *Reason, Truth, and History*, Cambridge: Cambridge University Press.
- Quine, W.V.O. (1960) *Word and Object*, Cambridge MA: MIT Press.
- Railton, P. (1980) *Explaining Explanation*, Ph. D. dissertation, Princeton University.

Salmon, W. (1990) *Four Decades of Scientific Explanation*, Minneapolis: University of Minnesota Press.

Salmon, W. (1994) Causality without Counterfactuals, *Philosophy of Science*, **61**, pp. 297-312.

Shannon, C. E. and Weaver, W. (1949) *The Mathematical Theory of Communication*, Urbana: Univ of Illinois Press.

Siewert, C.P. (1998) *The Significance of Consciousness*. Princeton: Princeton University Press.

Tye, M. (1994) Qualia, Content, and the Inverted Spectrum, *Noûs*, **28**, pp. 159-83.