

# **Philosophy's Movement Toward Cognitive Science**

By Dr. Charles Wallis

Last revision: 4/7/2013

## **Chapter 6 Identity Theories**

### **6.1 Introduction**

### **6.2 Identity Theories: Type-Type Identity**

#### **6.2.1 Two Arguments For Type-Type Identity Theory**

### **6.3 Identity Theories: Token-Token Identity**

### **6.1 Introduction**

The last chapter and lecture traces the developments that lead to the semantic twist. Specifically, chapter five and the associated lectures outline how philosophers of mind alter their focus from formulating and defending various ontological frameworks to understanding the relationship between scientific theories of the mind and ordinary conceptions of the mind. The shift in emphasis comes in reaction to the dramatic advances in science, mathematics, logic, and especially psychology at the dawning of the 20<sup>th</sup> century. The emergence of logical empiricism as the predominant philosophy of science strongly shapes how philosophers understand the engine that drives the rapid advances in the sciences at the beginning of the 20<sup>th</sup> century. Specifically, the logical empiricists focus upon scientific categorizations. The doctrine of the verification theory of meaning postulates that categorizations drive scientific progress only when they get their meaning through operationalizations—through being tied to empirical phenomena through intersubjective procedures and observations. Likewise, intertheoretic reduction and the unity of science emphasize semantic reduction as the key element of intertheoretic equivalence. As a result, philosophers begin increasingly to think of theories not simply in terms of theoretical posits, but in terms of the relationships between categories of scientific theories and the terms of ordinary language and of perceptual experiences.

After the semantic twist, therefore, most philosophers seek to understand how monistic physicalism can accommodate mental phenomena. Specifically, philosophers seek to advocate monistic physicalism by subsuming mental phenomena within the physical. Philosophers do not posit physicalistic theories providing mechanistic explanations of mental phenomena. Rather, philosophers seek to demonstrate the viability of monistic physicalism through semantic reduction. In other words, philosophers adopt the project of demonstrating how ordinary, pre-theoretic talk about the mind expresses nothing over and above what one can express using physicalistic language. As a result, philosophers adopt the methodology of demonstrating the viability of monistic physicalism by showing that one can specify the meanings of ordinary language mental terms completely using only physicalistic terms. Since philosophers do not formulate physicalistic theories of mental phenomena, they rely primarily upon scientific results to provide the data upon which they base their reductionist theories. The underlying inference driving such identifications lies in the notion that by directly equating mental and physical terms one indirectly identifies the referents of those terms—mental phenomena and physical phenomena. In other words, 20<sup>th</sup> century philosophers locate the explanatory

problem for the monistic physicalist in an inability to recognize or gather sufficient evidence for the co-referential nature of mental and physical terms.

The last chapter outlines one of two general approaches to pursuing this general strategy emerge within philosophy during the first half of the 20<sup>th</sup> century. On the one hand, theorists try to identify the meaning of mental terms with sets of overt, observable, physical behaviors definitive of those terms. Analytical behaviorism, discussed in the last chapter and lecture represents the first instance of this strategy. On the other hand, theorists seek to identify the reference of mental property, process, and entity terms with the reference of physical property, process, and entity terms through something less than, but sufficiently akin to analytical reduction. The current chapter and lectures explore theories within this second strategy, often called type-type reductionism. The delineation of identity theories in this lecture and chapter will emphasize two important features of type-type identity theories. First, identity theories operate within the general framework of logical empiricism. Specifically, identity theories view intertheoretic reduction as proceeding through semantic reduction of higher-level theoretical categories to lower-level physical categories. Second, while analytic behaviorism views reduction as requiring meaning equivalence, identity theories seek a weaker identification--specifically, reference equivalence. In other words, identity theories do not seek to understand intertheoretic reduction through meaning reduction. Instead, identity theorists suppose that intertheoretic reduction requires only co-reference. Thus, type-type identity theory asserts that folk mental concepts prove co-referential with physiological concepts. In contrast, token-token identity theorists only require that tokens of mental types are also tokens of some or other physical type.

## 6.2 Identity Theories: Type-Type Identity

Theorists often call the successor to logical behaviorism **type-type reductionism** or **type-type identity theory**. Type-type reductionism proposes to identify types of mental entities, mental properties, and mental processes with specific types of physical entities, physical properties, and physical processes. For instance, one might identify the mental property of pain with the physical property of stimulated c-fibers. Historians generally credit the British philosopher and psychologist U.T. Place (1924-2000) and the Austrian philosopher Herbert Feigl (1902-1988) as the source of the modern identity version of type-type physicalism. Place's colleague J.J.C. Smart (1920- ) also adopts this position. The motivations of identity theorists stem in large part from (and build upon) difficulties with logical behaviorism. For instance, Place tells readers,<sup>1</sup>

The view that there exists a separate class of events, mental events, which cannot be described in terms of the concepts employed by the physical sciences no longer, commands the universal and unquestioning acceptance amongst philosophers and psychologists which it once did. Modern physicalism, however, unlike the materialism of the seventeenth and eighteenth centuries, is behaviouristic. Consciousness on this view is either a special type of behaviour, 'sampling' or 'running-back-and-forth' behaviour as Tolman (1932,p. 206) has it, or a disposition to behave in a certain way, an itch for example being a temporary propensity to scratch. In the case of cognitive concepts like 'knowing', 'believing', 'understanding', 'remembering' and volitional concepts like 'wanting' and 'intending', there can be little doubt, I think, that an analysis in terms of dispositions to behave (Wittgenstein, 1953; Ryle, 1949) is fundamentally sound. On the other hand, there would seem to be an intractable residue of concepts clustering around the notions of consciousness, experience, sensation and mental imagery, where some sort of inner process story is unavoidable (Place, 1954). It is possible, of course, that a satisfactory behaviouristic account of this conceptual residuum will ultimately be found. For our present purposes, however, I shall assume that this cannot be done and

that statements about pains and twinges, about how things look, sound and feel, about things dreamed of or pictured in the mind's eye, are statements referring to events and processes which are in some sense private or internal to the individual of whom they are predicated. (p.44)

Two central ideas define type-type identity: **First**, Place and Feigl hold that behavioristic and identity analyses of mental terms do not exhaust the meaning of mental terms in ordinary language. That is, the new definitions of mental terms are not analytic—they do not capture the individually necessary and jointly sufficient conditions thought to dictate the meanings of ordinary terms. Place and Feigl hold that mental and physical terms pick out classes or kinds of things in virtue of their meanings, and that a significant part of the meaning of ordinary mental terms (as well as of the identity theorists' new analyses of those terms) is synthetic—i.e., going beyond the definitional meaning, usually as a result of experience. Specifically, type-type reductionists hold that the various behavioral associations between mental terms and physical/bodily terms serve to provide an initial description of a physical (brain) state. One can modify the initial behavioral descriptions, to the extent necessary, as a result of experience. Such descriptions ultimately determine the physical state that corresponds to the mental state. The identification of the physical state with the mental state constitutes a synthetic discovery. **Synthetic truths** or discoveries differ from analytical truths or discoveries in that analytical truths follow from the meaning of the terms involved. **Analytic truths** equate statements or terms and have truth-values dependent upon the meaning of the constitutive terms and/or logical structure. Denying an analytic statement therefore results in a contradiction. For example, "All triangles are three-sided planar figures," is an analytic statement, so is "A square is a four-sided, regular planar figure." Likewise, "It's false that all unmarried human males are bachelors" is a contradiction, since by definition bachelors are unmarried men. Analytic statements are said to be true by definition and/or their logical structure. In contrast, synthetic truths do not follow merely from the meaning of the terms involved. Someone can negate a synthetic statement without a contradiction. For example, one can assert the negation of the statement that "Long Beach has a population of 360,000 people," without a contradiction. "It's false that Long Beach has a population of 360,000 people." is not a contradiction. For this reason, discovering the population of Long Beach is a synthetic discovery.




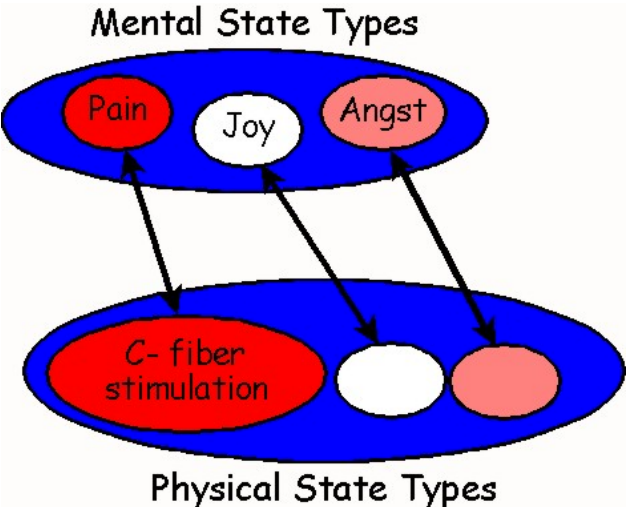

The synthetic nature of the discovery that brain states and processes prove identical to mental states and processes allows type-type theorists to side-step many of the objections raised by dualists. Place tells readers,

To say that statements about consciousness are statements about brain processes is manifestly false. This is shown (a) by the fact that you can describe your sensations and mental imagery without knowing anything about your brain processes or even that such things exist, (b) by the fact that statements about one's consciousness and statements about one's brain processes are verified in entirely different ways and (c) by the fact that there is nothing self-contradictory about the statement, 'X has a pain but there is nothing going on in his brain'. (p.45)

In other words, mental terms and physical terms share the same referent, but not the same meaning. Philosophers refer to terms that have the same referent, but which do not necessarily have the same meaning in all contexts as referentially opaque. Recall, two terms are referentially opaque if they refer to the same object or property, but one cannot intersubstitute the terms *salva veritate* (i.e. without changing the truth value of the statement) in many contexts—or instance belief descriptions. For example, people in northern latitudes can watch the northern lights (aurora borealis) regularly. They may proclaim that they see the

northern lights or that the northern lights are beautiful. However, most would not say that they are watching the photonic discharge resulting from ionized nitrogen atoms regaining an electron and nitrogen and oxygen atoms returning to a grounded state from an excited state after collisions with charged particles (solar winds) traveling along the magnetic field lines of the Earth’s magnetosphere. Nor would most observers likely say that the photonic discharge is beautiful. The difference, according to Place, between saying that the northern lights are photonic discharge and saying that the northern lights are beautiful is a logical feature:<sup>1</sup>

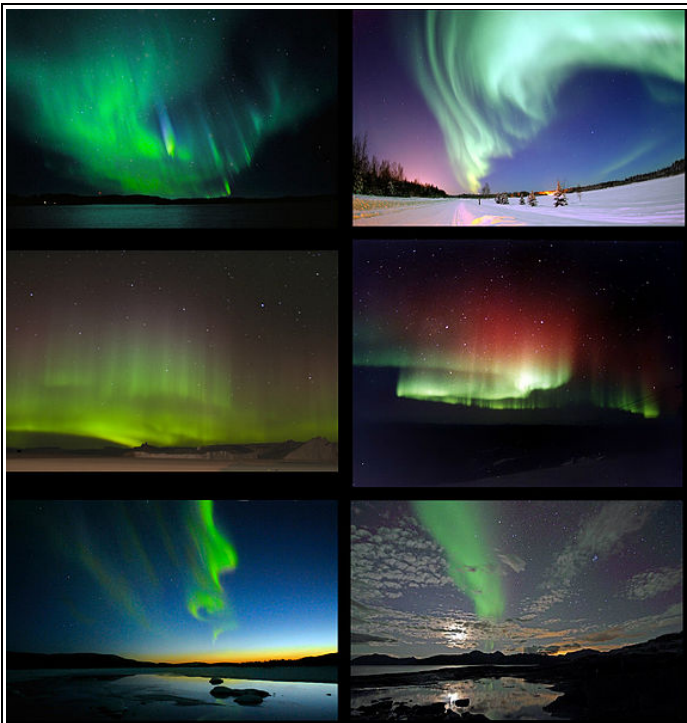
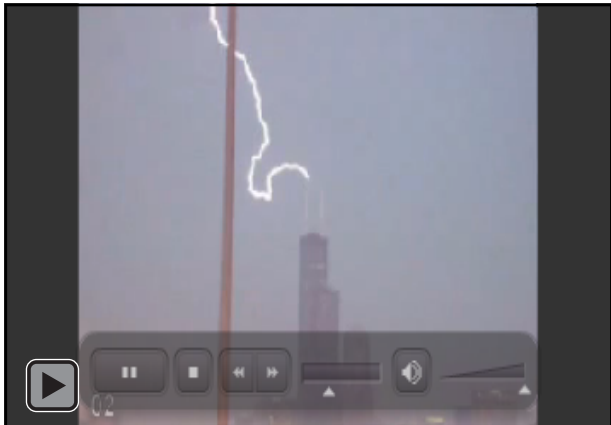
This logical feature may be described by saying that in both cases both the grammatical subject and the grammatical predicate are expressions which provide an adequate characterization of the state of affairs to which they both refer. (p.45)

		
Herbert Feigl (1902-1988) Modified From: <a href="#">University of Minnesota</a>	U.T. Place (1924-2000) Modified From: <a href="#">Adelaide University</a>	John Jamieson Carswell "Jack" Smart (1920- ) Modified From: <a href="#">Monash University</a>
		
Diagram depicting Type-Type Identity Theory. Type-type theorists claim that there is only physical substance. Mental properties exist, but are type-identical to types of physical properties. Theorists discover the identity using the physical and behavioral associations between mental and physical terms to identify the physical state type corresponding to the mental state type.	Pictures of U.T. Place’s brain located at the Anatomy Museum in Medical School Building at the University of Adelaide. Place taught as a lecturer in Philosophy and Psychology at Adelaide University from 1951-1954. From: <a href="#">Adelaide University</a>	

Feigl expresses a similar point by telling readers that:<sup>2</sup>

The identity thesis which I wish to clarify and to defend asserts that the states of direct experience which conscious human beings "live through," and those which we confidently ascribe to some of the higher animals, are identical with certain (presumably configurational) aspects of the neural processes

in those organisms. ...we may say, what is *had-in-experience*, and (in the case of human beings) knowable by acquaintance, is identical with the object of knowledge by description provided first by molar behavior theory and this is in turn identical with what the science of neurophysiology describes (or, rather, will describe when sufficient progress has been achieved) as processes in the central nervous system, perhaps especially in the cerebral cortex. In its basic core this is the "double knowledge" theory held by many modern monistic critical realists. ... The "mental" states or events (in the sense of raw feels) are the referents (the denotata) of the phenomenal terms of the language of introspection, as well as of certain terms of the neuro[-]physiological language. For this reason I have in previous publications called my view a "double-language theory." But, as I have explained above, this way of phrasing it is possibly misleading in that it suggests a purely analytic (logical) translatability between the statements in the two languages. It may therefore be wiser to speak instead of *twofold access* or *double knowledge*. The identification, I have emphasized, is to be empirically justified, and hence there can be no logical equivalence between the concepts (or statements) in the two languages. (Section E ¶4 and 5)

	<p>(Right) Several pictures depicting the beautiful photonic discharge resulting from ionized nitrogen regaining atoms electron and nitrogen and oxygen atoms returning to a grounded state from an excited state resulting from their collisions with charged particles (solar winds) traveling along the magnetic field lines of the Earth's magnetosphere. Er, the beauty of the northern lights (aurora borealis). From: <a href="https://en.wikipedia.org/wiki/Aurora_borealis">Wikipedia</a></p>
	<p>(Below) A movie of the sudden large-scale discharge of electrons between bodies of suspended liquid and frozen crystals measuring no more than a few tens of microns and having a negligible fall velocity, i.e., a lightening discharge between clouds. Click to play. From: <a href="https://www.youtube.com/watch?v=Q2">Youtube</a></p>
	

### 6.2.1 Two Arguments For Type-Type Identity Theory

Place and Feigl, then, have two lines of argument for their identity theory. On the one hand, they argue that traditional dualist arguments like the argument from Leibniz's Law and the argument from introspection do not apply to the identities Place and Feigl advocate. Call these the **inapplicability arguments**. On the other hand, Place and Feigl present an argument from analogy with other historical examples of scientific reduction. Call these the **reductive analogy arguments**. Object color is really surface reflectance. Heat is mean kinetic energy. Sound is compression waves traveling through the atmosphere and pitch is just the oscillatory frequency of those waves. Lightening is just the sudden large-scale discharge of electrons between clouds. Feigl tells readers that:<sup>2</sup>

...the advance of scientific theories consists essentially in the reduction of a variety of originally heterogeneous observable facts and regularities to a unitary set of explanatory concepts and postulates. Customarily it is said, for example, that visible light is electromagnetic radiation (within a



certain interval of wave lengths); that table salt is NaCl; that magnetized iron is an aggregate of iron atoms with a characteristic spin of certain of their electrons; that the transmitters of hereditary traits are the genes in the chromosomes of the germ cells; that (at least) short range memory traces are reverberating circuits in cerebral cell assemblies, etc. The "is" and the "are" in these sentences represent identities. But these identities differ in their mode of certification from the analytic identities of pure logic and mathematics. ... ..the identities established in the factual sciences are confirmed on the basis of empirical evidence. ... ..there are also such empirically ascertainable identities as those of Tully and Cicero, of William Thompson and Lord Kelvin, or of the evening star and the morning star. In the examples just given we have (extensional) identities of individuals labeled or uniquely described in two or more ways. (Section D, ¶2 &3)

### 6.3 Identity Theories: Token-Token Identity

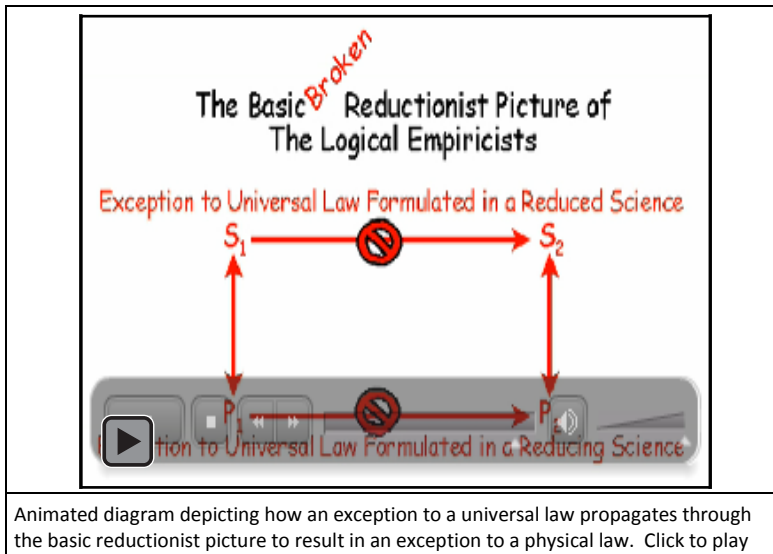
In 1970, Donald Davidson (1917-2003) proposes a new version of identity physicalism. Davidson starts his chapter, "[Mental Events](#)",<sup>3, 4</sup> by stating his motivation for the view,

Mental events such as perceivings, rememberings, decisions, and actions resist capture in the nomological net of physical theory. .... I start from the assumption that both the causal dependence and the anomalousness of mental events are undeniable facts. My aim is therefore to explain, in the face of apparent difficulties, how this can be. (p.138)

By **nomological** Davidson simply means law-like, and by **anomalousness** Davidson means not falling under exceptionless universal laws. Davidson thinks that the failure to find the simple type-type identifications theorists like Smart, Place, and Fiegl suggest together with the failure of psychology and sociology to generate universal exceptionless laws warrants a reconsideration of the type-type identity theory. In the quote above, Davidson tells readers that he takes the anomalousness of the mental as an undeniable fact. That is, Davidson holds that psychologists have not and cannot discover universal, exceptionless laws formulated in mentalistic categories.

The anomalousness of the mental presents a problem for Davidson because he accepts both (1) the monistic physicalistic framework and (2) the basic logical empiricist picture of laws and reduction in science. That is, Davidson holds that the dramatic success of sciences like physics proves the mechanistic and deterministic nature of the physical world as one describes it using the physical conceptual scheme. Specifically, scientists formulate laws using exclusively physicalistic descriptions. These physicalistic descriptions represent a conceptual scheme for describing the world. This conceptual scheme has proven itself capable of describing the world so that the sciences produce finite, exceptionless universal laws. Indeed, the collection of such laws forms a closed, complete deductive system. That is, given a complete physicalistic description of some state of the world, called a physical event, scientists can, at least in principle, deduce how the world will unfold by deducing the resulting physical event, i.e., the exclusively physicalistic description of the world resulting from the prior event.

The standard view of physical laws looks like the bottom of the diagram (below) for the reductionist view of science. However, what happens to the picture if the mental has exceptions to its laws? If one supposes that physical laws and bridge laws between the mental terms and the physical terms are exceptionless and universal, then an exception to a psychological law between  $S_1$  and  $S_2$  is an exception to the physical law between  $P_1$  and  $P_2$ . That is, all members of  $S_1$  are members of  $P_1$  and all members of  $S_2$  are members of  $P_2$ .



Thus, the exception to the psychological law between a token (member) of  $S_1$  and token (member) of  $S_2$  is also an exception to the physical law. The exceptionless bridge laws mean that the token (member) of  $S_1$  is also a token (member) of  $P_1$  and token (member) of  $S_2$  is also a token member of  $P_2$ . Thus,  $P_1$  and  $P_2$  violate the physical law just as  $S_1$  and  $S_2$  violate the psychological law.

Yet, Davidson famously argues that the anomalousness of the mental proves consistent with a certain sort of physicalism in his paper.

However, many people think that Davidson offers no reason to accept the anomalousness of the mental. In fact, Davidson does offer a reason for the anomalousness of the mental. Mental terms, like physical terms, form a conceptual scheme. Unfortunately, the conceptual scheme for the mental differs from the conceptual scheme for the physical. Davidson holds that the physical conceptual scheme has as its exclusive purpose conceptualization of the physical world for the purpose of formulating and testing physical laws. In contrast, the primary purpose of the mental conceptual scheme consists in making attributions of mental terms so as to provide one with an understanding of other people. Specifically, Davidson argues that in order to understand the actions, beliefs, desires, etc. of a person one must understand the person as a rational agent. In other words, in so far as one cannot see the actions, thoughts, etc. of an individual as rational one cannot understand them. That is, the actions of the individual make no sense in the mental conceptual scheme if one cannot understand that person as acting in accordance with one's own standards of reasonableness. This emphasis on rational understanding can, and Davidson suggests does, often trump ascriptions that would support universal and exceptionless laws in psychology. Davidson feels that one can use psychological generalizations to explain and predict by relating mental events, i.e., descriptions of some state of the world using exclusively mentalistic terms. However, mental events, events one describes using mental terms, do not lend themselves to the expression of exceptionless universal mental laws because the mental conceptual scheme works to maximize one's ability to understand others, even if such understanding violates psychological laws. Hence, psychological laws, as mere heuristic generalizations, do not form a closed, complete deductive system.

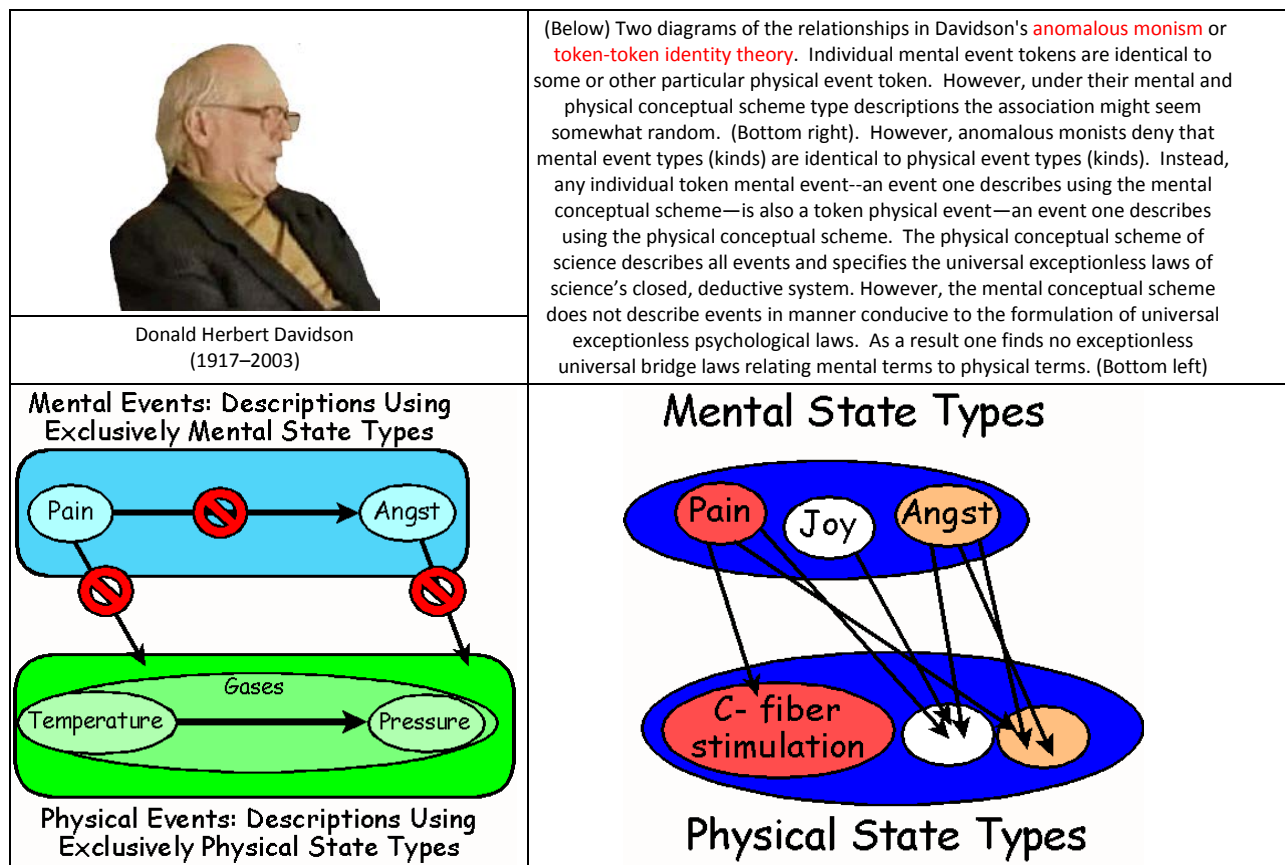
But how can one avoid the seeming dire consequences of an anomalous mental realm? Specifically, Davidson wants to preserve three principles:<sup>3</sup>

(1) Causal interaction occurs between the mental and the physical. (pp.137-138)

(2) If causal interaction occurs, then strict deterministic laws (universal and exceptionless) govern that interaction. Davidson refers to this property as "the Principle of the Nomological Character of Causality." (p.138)

(3) “The third principle is that there are no strict deterministic laws on the basis of which mental events can be predicted and explained (the Anomalism of the Mental).” (p.138)

How does Davidson solve his difficulty of rendering these principles consistent? The answer lies in his notion of the two conceptual schemes. Davidson holds that since the mental and physical conceptual schemes differ in their goals, no possible reduction of mental terms to physical terms exists. In other words, Davidson denies the possibility of exceptionless universal "bridge-laws"--laws relating mental descriptions of states of the world (mental events) to physical descriptions of states of the world (physical events). Though the argument seems complex, it's actually pretty straightforward. Physical laws are universal and exceptionless. Mental laws are neither universal nor exceptionless. As a result, if there were universal and exceptionless laws linking mental and physical events (bridge-laws), then those bridge-laws would provide a basis for universal exceptionless mental laws. Yet, there are no exceptionless universal mental laws. Conversely, if there were universal and exceptionless laws linking mental and physical events (bridge-laws) and mental laws has exceptions, then those universal and exceptionless bridge laws would transfer those exceptions of mental laws into exceptions in physical laws. Yet, there are no exceptions to physical laws. Hence, there can be no universal and exceptionless bridge laws.



What about the causation between the mental and the physical (principle 1) and the strict deterministic nature of all such causation (principle 2)? Easy, says Davidson; mental events are just descriptions of the world using the mental conceptual scheme. Every mental event is just an event described using mental terms. But one can also describe that event using physical terms from the physical conceptual scheme. Thus, every token of a mental event is also a token of a physical event. The physical and mental descriptions might only hold for that token event, or the descriptions may prove more general. However, mental types do not reduce



to physical types. So, some tokens of a particular type of mental event, call it  $M_1$ , will have token physical descriptions from different physical types than other tokens of mental events from that type, .i.e.  $M_1$ . Thus, despite the lack of mental-type to physical-type reduction, one can understand how token mental events are identical to token physical events—they are just different descriptions of the same event.

Moreover, since every token of a mental event has a physical event description, token mental events can causally interact with token physical events in a strict deterministic manner—the manner dictated by the strict deterministic law relating their physical event descriptions. Since physical event descriptions yield finite, exceptionless universal laws that combine to form a closed deductive system, all tokens of mental causal interaction with the physical fall under strict deterministic laws—just not laws using mental terms.

### **Bibliography**

1. Place UT. Is Consciousness a Brain Process? *British Journal of Psychology*. 1956;47:44-50.
2. Feigl H. The 'Mental' and the 'Physical'. In: Feigl H, Scriven M, Maxwell G, eds. *Concepts, Theories and the Mind-Body Problem, Minneapolis, Minnesota Studies in the Philosophy of Science*. Vol 2. Minneapolis, MN: University of Minnesota Press; 1958.
3. Davidson D. Mental Events. In: Vermazen B, Hintikka MB, eds. *Essays on Actions and Events*. Oxford: Oxford University Press; 1980.
4. Davidson D. Mental Events. In: Foster L, Swanson JW, eds. *Experience and Theory*. London: Duckworth; 1970.