

# On the Nature of Relations Learned in Pavlovian Conditioning

Víctor García-Hoz Rosales<sup>1</sup>

*Universidad Complutense de Madrid, España*

## ABSTRACT

The statement is frequently made that in Pavlovian conditioning the animal learns or acquires the knowledge of a causal relation between the conditioned stimulus and the reinforcer. Two points can be distinguished in the statement: (1) the animal learns that there is a relation between the stimuli, and (2) the relation the animal learns is a causal relation. Leaving aside the first question, the second point -i.e., the type of relationship learned by the animal- may be approached by asking what kind of relation there exists between the conditioned stimulus and the reinforcer in the Pavlovian procedure. If we suppose that the animal only learns about those relations that he in fact experiments, then the relation between stimuli learned by the animal in Pavlovian conditioning will be bounded by the relation between stimuli prevailing in conditioning. In this way, if Pavlovian conditioning allows the animal to learn a causal relation between the stimuli this would be first and foremost due to the effective occurrence of such a causal relationship. And it can be pointed that it is not easy to describe the stimulus-stimulus relationship operating in Pavlovian conditioning as a causal one, seeming much more appropriate to characterize it as a signal relation. From this point of view, and if one wishes to preserve the notion of learning of relations, it would be wiser to say that in Pavlovian conditioning the animal learns the relation between a signal and the thing it signals, which is not a the relation between cause and effect.

*Keywords:* Causal and Signal Relations, Content of Learning.

## RESUMEN

*Sobre la naturaleza de las relaciones aprendidas en el condicionamiento pavloviano.* Es común afirmar que, en el condicionamiento pavloviano, los animales adquieren o aprenden la relación de causalidad entre el estímulo condicionado y el reforzador. Dos partes se pueden distinguir en la afirmación: que los animales aprenden que hay relación entre los estímulos y, además, que la relación que aprenden es una relación de causalidad. La clase de relación que el animal aprende puede abordarse examinando qué clase de relación existe entre el estímulo condicionado y el reforzador en el procedimiento de Pavlov. Si suponemos que el animal no aprende otras relaciones que las que experimenta, la relación entre estímulos que el animal pueda aprender en el condicionamiento vendrá limitada por la relación entre estímulos que en el condicionamiento exista. De este modo, si con el procedimiento de Pavlov el animal aprende relaciones de causalidad entre los estímulos será, primero y principal, porque en esa situación se establezca entre los estímulos una relación de causalidad. No parece fácil describir la relación estímulo-estímulo que caracteriza el condicionamiento pavloviano como una relación de causalidad, sino que resulta más plausible entenderla como una relación de señal. Desde este punto de vista, y si se quiere mantener la noción

<sup>1</sup>An early version of this paper was presented at the joint SEPC\ISPC Meeting held in Oviedo, Spain, September 2004.

Correspondence concerning this article should be addressed to the author: Facultad de Psicología, Universidad Complutense de Madrid, Campus de Somosaguas, 28223 Madrid, España. E-mail: vghoz@psi.ucm.es.

de aprendizaje de relaciones, quizá fuera más apropiado decir –con algunos autores ya lejanos en el tiempo- que en el condicionamiento pavloviano el animal aprende la relación entre una señal y la cosa señalada, que no es la relación entre la causa y el efecto.

*Palabras clave:* relaciones causales y de señal, contenido del aprendizaje.

Several modern authors (Dickinson, 1980; Mackintosh, 1977; Rescorla & Holland, 1976; Testa, 1974) seem to agree in the description or conceptualization of Pavlovian conditioning as a kind of experience by which an animal learns or establishes causal relations between stimuli. There are two parts to this conceptualization: On the one hand, that the animal establishes some relation between the stimuli and, on the other hand, that the relation thus established is a causal one. Leaving aside the first part, that is, taking for granted that in Pavlovian conditioning the animal learns about a relation between stimuli, the second part raises the question about the type of relation learned or established. The following words are to be understood as an attempt to draw attention to a particular aspect of this issue, namely, the kind of relation between stimuli effectively occurring in the Pavlovian conditioning situation.

### RELATIONS LEARNED IN PAVLOVIAN CONDITIONING

Within the framework of a cognitive concept of conditioning, many modern writers have adopted what, for lack of a better term, could be called a causal relation hypothesis. According to this hypothesis, in Pavlovian conditioning, animals establish a relation of causality between the conditioned stimulus and the unconditioned stimulus or reinforcer. By contrast, previous authors of various conceptual outlooks had opted for a different point of view, describing conditioning as the establishment of a signal relation between the conditioned and unconditioned stimuli.

#### *The Causal Relation Hypothesis*

In Rescorla & Holland (1976) appears what can be called the emblematic expression of the causal hypothesis, which not only includes this hypothesis but also the non-cognitive hypothesis that Rescorla and Holland considered to be the opposite. Claiming to belong to the “psychological associationism tradition,” which opposes the “physiological tradition of reflexology,” they wrote that: “[W]e view conditioning (more) as the way in which the organism learns about the causal relations in his environment (what Tolman & Brunswik [1935] called the «causal texture») than as the transfer of control of a reflex from one stimulus to another” (p.172).

In a brief review of the causal hypothesis, it can be mentioned that its first modern formulation is to be found in the work of Testa (1974), whom subsequent authors have quoted as a source of inspiration. Testa argued that the stimuli that the animal experiences in a conditioning situation are related through a mechanical causality and that the conditioning procedures could therefore be described as an “exposed causal relationship” and the association as an “internalized causal relationship” (pp.493-495).

Mackintosh (1977), on the other hand, considers the effect of contingency on the

selectivity of the stimulus in Pavlovian conditioning to conclude that: "By conditioning selectively in the way we have seen, laboratory rats succeed in attributing the occurrence of reinforcers to their most probable causes" (p.247). In other words: "We know (...) that two events will be associated only to the extent that the first is a better predictor of the second than any other event is, and thus that associative learning is nicely designed to allow animals to keep track of the causal structure of their world" (Mackintosh, 1997, p.881). Lastly, it can be noted that in his influential *Contemporary Animal Learning Theory*, Dickinson (1980) embraced the causal hypothesis and turned it into a distinctive trait of this "contemporary animal learning theory." As though to eliminate any trace of a doubt when describing Pavlovian conditioning, Dickinson avoided the terms *conditioned stimulus* and *unconditioned stimulus*, using instead those of *cause* and *effect*, respectively: "In the typical classical conditioning experiment, E1, the potential cause, is usually a neutral stimulus and E2, the potential effect, is typically a motivationally significant stimulus" (p.19). In Pavlovian conditioning, Dickinson insisted, the animal is "observing causal chains" (p.19).

The defenders of the causal hypothesis usually add on an adaptive argument. They say that it is beneficial, perhaps even necessary for the animal's survival, to learn the causes of events that occur in the environment. ("So it appears that for an animal to act adaptively, it must be capable of detecting and storing information about the causal structure of its environment." Dickinson, 1980, p.8.) And this is precisely the function that is postulated for conditioning: "Associative learning mechanisms have been shaped by evolution to enable animals to detect and store information about real causal relationships in their environment" (Dickinson, 1980, p.26). Thus, by describing it as the way an animal learns causal relations, an obvious adaptive or biological significance is assigned to Pavlovian conditioning, which might not be so clear -one could surmise- in an alternative conceptualization of conditioning.

In any case, the causal relation hypothesis is not much more than the previous statements and it does not constitute -nor does it intend to- an explanatory theory, but rather a generic conceptualization or, if preferred, a point of view: "The view that conditioning is a primitive form of causal analysis [is not offered] as an *explanation* of any of the facts of conditioning; I do believe, however, that it provides an equally valid and rather more interesting way of thinking about conditioning than that offered by traditional S-R theory." (Mackintosh, 1977, p.247, original italics). In any case, it may be of speculative relevance or, at least, a not entirely useless exercise to examine in more detail -at that level of point of view or general characterization- the notion of causal acquisition and, in particular, to compare it with what could be called the signal relation hypothesis.

### *The Signal Relation Alternative*

In effect, the proposal of the acquisition of a causal relation is at variance with the one that is linked to some classic authors, such as Pavlov and Tolman, and according to which the relations that animals establish in Pavlovian conditioning are signal relations, where the conditioned stimulus is the signal and the unconditioned stimulus is the thing

signaled. Despite their different orientation and theoretical significance, Pavlov and Tolman share the hypothesis that, in Pavlovian conditioning, animals learn or establish signal relations or, in other words, that the conditioned stimulus becomes the signal of the unconditioned stimulus, which is thus signaled. It is to be reminded that although Pavlov did not use a learning of relations terminology, he stated expressly and repeatedly that conditioning is a “signaling activity” by which the conditioned stimulus becomes the signal of the unconditioned stimulus, and that the conditioned reflex should be considered a “signal reflex” (1927, cap. 1). Nor is it pointless to remember that Tolman (1932, chap. 21) described Pavlovian conditioning as the acquisition of a “sign Gestalt” (more exactly, a “Gestalt-sign-expectation”), made up of three elements or parts: the conditioned stimulus as a sign object, the unconditioned stimulus as a significate, and the signified means-end relation. Lastly, “the conditioned response is a response to the sign-relationship  $S_2 \rightarrow S_1$ ” (1932, p.330). It is worth noting that Tolman is a recognized antecedent of the cognitive approach to the study of conditioning, from whose framework the causal hypothesis seems to emerge.

It is remarkable that the authors who resort to the causal relation should avoid contrasting their outlook with the classic conceptualization that appeals to the signal. These authors express no criticism of the concept of signal applied to conditioning, and there is no analysis which reveals the shortcomings of that concept and the need to substitute it with the concept of cause; or, if not strictly a need, the advantages of so doing. (More understandably, neither do the authors who refer to Pavlovian conditioning as “signaling” consider the possibility -in fact, they don’t even mention it- of applying causal notions to it.) In a word, it is unclear why the concept of cause is preferred to and replaces the concept of signal; it is, in fact, amazing. For instance, it seems evident that the covariation conditions that, according to Mackintosh (1977, 1997), allow the animal to determine the cause of a certain effect are no different from those that allow it to determine the signal of a certain event. But Mackintosh insists on talking about causal relation and does not even mention the possibility that, precisely because of the covariation, a signal relation is established.

As it can be gathered from some of the preceding quotations, the acquisition of causal relations is proposed by way of opposition to the so-called “traditional S-R theory,” according to which the only thing an animal acquires in conditioning is an stimulus-response connection, a movement connected to a stimulus, not the knowledge of a relationship. Note, however, that for the purpose of contrasting learning a movement with learning a relationship, the concept of signal could have been just as useful as that of cause. Perhaps the concept of cause is preferred because the concept of a signal - for what it’s worth- is used and fits in with the Pavlovian theory, which is not considered unrelated to the “traditional S-R theory.” Ultimately, it seems that the causal proposal depends, first and foremost, on the wish -whether or not on target- to set straight a cognitive viewpoint of conditioning.

### *An Aspect of the Issue*

No matter what, one way to address the issue that divides the two viewpoints, that is, the type of relation that the animal learns in Pavlovian conditioning, is to

examine the type of relation between the stimuli that objectively occurs in that conditioning situation. This issue -and, in fact, the intention of this work and the distinction between the two conceptualizations- is valid if we start with the assumption that, among others, two different types of relations can occur between two events: a relation of cause and effect and a relation of signal and signaled thing. Assuming this, it makes sense to ask which of these relations between stimuli -signal or causality- actually occurs in the learning situation known as Pavlovian conditioning. Because if it is true that the environmental events determine the animal's experience and the resulting learning, so that the animal acquires a representation of its environment, then the nature of the objective relation between the stimuli should restrict the nature of the relation that the animal establishes. If there is a relation A and a relation B, and an animal experiences the relation A, the animal may or may not learn this relation A, but it would be difficult to say that it had learned a relation B. In other words, it seems clear that, to learn a relation, the animal has to experience it. Therefore, in order to determine the nature of the relation the animal learns in Pavlovian conditioning, it does not seem pointless -it is perhaps unavoidable- to wonder what relation actually exists between the stimuli in this situation, because that is the relation that the animal will experience and be able to learn. This particular question about the nature of the relation between stimuli that is established in Pavlovian conditioning was the origin of this work.

### CONCEPTS OF SIGNAL AND CAUSE

At this point, and before going on, it is necessary to halt briefly at the notions of causal relation and signal relation or, ultimately, cause and signal. Armed with some knowledge of the relations between events in the world and the meaning of words, anyone can distinguish between a cause and a signal, the same as between the relation that links an event with something that proclaims it and the relation that links an event with something that produces it. It is easy to understand that there is a difference between covariation and causation and, ultimately, that the statement "*post hoc, ergo propter hoc*" ("After this, therefore because of this") is a fallacy. Hence, to elucidate the nature of the Pavlovian relation, one could resort to some kind of example and search for parallelisms with Pavlov's procedure. However, with a view to the following, it might be preferable to define the concepts of signal and cause. Of course, this does not intend to solve the difficult problems that the concept of cause, for example, entails. The purpose is simply to clarify or explore the concepts beyond their mere denomination and, in any case, it is understood that the proposed definitions are representative of generally accepted viewpoints of both concepts.

#### *Signal*

Signal is usually defined, first, in terms of the mental or psychological operation it produces; that is, in terms of the action of what is called a signal on a mind or a certain representation system. Thus, an old definition from the Scholastics says that signal is what "*per quod aliquid devenit in cognitionem alterius.*" The expression can be transcribed -and accepted as a definition- by stating that a signal is the object or

event that brings to mind the image or idea of another object or event. In somewhat different terms, but with the same functional meaning, a signal is the stimulus A that produces an internal or mental operation of evocation or recovery of a second stimulus B, and more simply, a signal is the stimulus A that activates the internal representation of another stimulus B. Hence, the signal represents the signaled insofar as the signal causes the mental operation originally belonging to what is signaled (to activate the internal representation of the signaled event). The operation of the signal can also be described in terms of behavior. It can be stated without substantially contradicting the former statements that a signal is “something that directs behavior with respect to something that is not at the moment a stimulus” (Morris, 1946, p.366).

From the viewpoint of the mental or internal operation, one could say that a signal is a stimulus whose occurrence has the property of activating its own representation and the representation of another stimulus or, which seems to amount to the same thing, a stimulus whose representation, when excited, activates in turn the representation of a second stimulus in the absence of the second stimulus. In the cases in which it activates a representation of what is signaled prior to its occurrence, one could say that the signal indicates the existence of a future event. And in that case, one could think that the signaling mechanism’s function, perhaps first and foremost, is to announce or predict the occurrence of the signaled event.

Having said that a signal relation is the mental or internal relation that exists between two stimuli, one of which refers to the other in the animal’s mind or behavioral mechanism, the following question could be posed -the concept of signal would be incomplete if it is not answered- about why one stimulus becomes the signal of another stimulus. This question is commonly answered by assuming that a stimulus turns into the signal of another stimulus when they are in some physical and objective relation that the animal experiences and processes. That is, the signal has an internal relation with the idea or representation of what is signaled because in the animal’s environment (and experienced by the animal) there is an objective relation between the signal and the signaled. This being so, it has to be decided to which physical relations between the objects does the signal relation correspond, that is, under what circumstances does one of the events become the signal for the other when both events occur in the environment and are experienced by the animal.

At this point, the classic proposition, which essentially stands unrefuted, is that, in the simplest case, an object or event A becomes the signal for another event B, when A coincides with B in space and precedes B in time. More extensively, when two stimuli occur contiguously in time and space, and one precedes the other temporally, and such stimuli are experienced and elaborated by the animal, the subsequent occurrence of the first stimulus will lead to the animal’s internal operation of evocation or recovery of the second stimulus. Which amounts to say that the physical and external relation represented or recovered by the internal signal relation is a spatial-temporal contiguity relation, or in other words, that the spatial-temporal contiguity relation is the physical relation, experienced by the animal, which leads to an internal signal relation. (The term *contiguity* can be replaced by that of *covariation* without affecting the essence of the argument).

### *Cause*

The concept of cause denotes something different from a mental or knowledge operation and it refers to the property attributed to events of a very different physical nature of producing or changing another event from the physical universe, which is called the effect. It is clearly not psychology's task to ask why an event should have that property of influence over another event, nor to attempt to guess what this capacity for action consists of. It may be an epistemological question and belonging to different specific sciences but it is not a psychological question (except, obviously, for what refers to psychological causation).

The issue of the cause becomes psychological, however, when the question refers to the conditions in which a subject acquires knowledge of a causal relation. It is the so-called question of causal cognition, which the British empiricist David Hume studied with particular determination and resonance. And to answer the question of causal cognition, it is first necessary to determine the characteristics of a causality relation, in order to know whether, in a given situation, this relation actually occurs and the subject can, by experiencing it, learn it.

For the purposes of this article, we must consider the case of two discrete and discriminable facts from the environment -two stimuli like the ones employed in Pavlovian conditioning- and state precisely under what conditions the first stimulus is understood to be the cause of the second, that is, when can be said that the two stimuli are in a causal relation. Well then, according to a long tradition (Russell, 1945), it is understood that an event A is the cause of an event B when: (a) A and B coincide in time and space, (b) A precedes B, and (c) A exerts an action of physical production or alteration on B. These circumstances -and no others- are the circumstances that characterize a causal relation between two events occurring in the environment and therefore these circumstances -and no others- are those that, experienced by the animal, should allow it to establish a causal inference.

It is to be noted that the conditions of spatial-temporal contiguity and precedence that first define a causal relation are sufficient to establish a signal. But causality adds one defining and exclusive property to contiguity and precedence, and it is this property which ultimately distinguishes causality from signaling. Hume writes: "Shall we then rest contented with these two relations of contiguity and succession, as affording a complet idea of causation? By no means. An object may be contiguous or prior to another, without being considered as its cause" (1739/1978, Bk. I, Pt. III, Sect. II). In order to establish a cause, contiguity and precedence are necessary but not sufficient conditions. A cause is also defined by having the property of acting on the occurrence itself or on the occurrence characteristics of the event that follows it. So, a cause is a signal but with an addition that the signal, by definition, lacks. From this point of view, it seems clear that causality and signaling are different relations, whose learning or knowledge should correspond to different experiences and different ways of processing.

On the other hand, as it is just said, the cause shares with the signal the circumstances of spatial-temporal contiguity and precedence, from which follows that when a stimulus is represented as the cause of an effect, it is also represented as the

signal of that effect and, as such, it activates the internal representation of the effect before its occurrence. In other words, it is evident that a cause is (or can be), in addition to the cause, the signal for the effect it produces. Thus, the same announcing or predictive function can be and has been attributed to the cause as it is attributed to the signal. Sometimes, this functional coincidence -which can be easily accepted- has led, if not to identifying, at least to confusing cause and signal: given that a cause is a signal, it seems that a signal should be a cause; given that cause and signal both predict, they are the same thing.

But in no way the signal relation, based on contiguity, can be identified with the causal relation, based on physical influence, and, although it is true that every cause is the signal of its effect, not every signal of an event is its cause. Hence, it is possible to state that, concerning prediction, the cause is a kind of signal, but it does not follow that the signal is a kind of cause. In any case, it would be acceptable to treat the cause (also) as a signal, which it is, but not to treat a simple signal as a cause, because it may or not be.

Perhaps in order to avoid confusion, and when convenient, it could be distinguished between noncausal and causal signaling or between noncausal and causal prediction, attributing the former to a pure or noncausal signal, and the latter to a causal signal (cause proper). One could distinguish them not only for the sake of expository clarity, but also because it is conceivable that the prediction would be different in each case, because (hypothetically) causal prediction could be affected by the processing or knowledge of that added relation, which does not exist in pure prediction. Anyhow, the causal relation (which implies signaling) and the pure signal relation (which does not imply causality), although they share common properties, are effectively different and mutually excluding relations, and one cannot simultaneously sustain both relations about the same link between two events.

### SIGNAL AND CAUSE IN PAVLOVIAN CONDITIONING

The question is which relational circumstances occur between two stimuli in Pavlovian conditioning: Those that define a signal relation or those that define a causal relation? Considering the first stimulus: Is the conditioned stimulus experienced in circumstances in which it is presumed that one event becomes the signal of another event or it is rather experienced in circumstances in which one event is said to be the cause of the other?

#### *Pavlovian Conditioning as a Signal Relation*

To start with, it is obvious that Pavlovian conditioning, insofar as it is an operation with stimuli, is defined in terms of spatial-temporal contiguity and precedence between these stimuli. Thus, according to Pavlov, contiguity is the main condition to establish "new conditioned reflexes or new nervous connections: The fundamental requisite is that any external stimulus which is to become the signal in a conditioned reflex must

overlap in point of time with the action of an unconditioned stimulus.” (1927, p.26). And concerning precedence, Pavlov adds: “[I]t is also and equally necessary that the conditioned stimulus should begin to operate before the unconditioned stimulus comes into action.” (1927, p.27). It does not seem that, in its essential terms, one can question this description of the ideal conditioning circumstances, nor that the notions employed in this description can be substituted by others. From the point of view of experimental practice, the instructions for a standard Pavlovian experiment are that the conditioned stimulus be followed by the reinforcer, and the description of the procedure in this kind of experiment always includes the circumstances of the relation of contiguity and precedence between the stimuli (affirming or denying it, depending on the case), making no statement about its causal relation. Summing up, temporal contiguity and precedence are the characteristics that define the relation between stimuli in Pavlov’s experimental situation. That being so, and according to the notion of signal that we started out from, it can be stated that the relation established between stimuli in Pavlovian conditioning is the kind of relation that makes one stimulus to become the signal of the other.

It remains to be seen, however, whether in Pavlovian conditioning there is some kind of physical action or influence between the stimuli that would allow us to classify it as a causal relation situation. To be brief, it seems there is not. On the one side, the abstract definition of conditioning does not require a relation of physical action or influence of one stimulus over the other; in addition the description of the Pavlovian procedures does not include any mention of the establishment or demonstration of this relation of causality, in contrast to the relation of contiguity. In fact, it seems impossible to expressly postulate for any conditioning situation that the conditioned stimulus causes the reinforcer: It is evident that the bell that Pavlov’s dog heard was not the cause of its being fed shortly afterwards, nor in conditioned suppression, that the sound is the cause of the shock; just as in taste aversion, the taste is not the cause of the nausea that acts like an unconditioned stimulus. Experimentally the conditioned stimulus and the unconditioned stimulus are events that are programmed and that occur independently in all ways except for contiguity and precedence. Bringing to mind the Cartesians’ clocks (Russell, 1945) they are like two clocks that are set with a slight difference, so that one will ring before the other one, but that does not mean that the ringing of the first one causes the second one to ring.

However, perhaps one could refer to Pavlovian conditioning with respect not only to a situation of mere spatial-temporal contiguity but also to a situation of causality, as this latter situation includes the former. As mentioned above, just as the causal relation does not eliminate the signal relation, neither does it destroy the Pavlovian nature of a situation, which in principle could be either a relation of pure signaling or mere temporal contiguity, or of causal signaling. But the question proposed herein does not refer so much to relations that can be included or added on to Pavlovian conditioning, as to the defining condition of the relation therein established. According to the above, one could postulate that this defining condition -necessary and sufficient of itself- is the very same spatial-temporal contiguity of the stimuli that defines the signal relation, without including the physical influence of one stimulus over another which is what characterizes the causal relationship.

### *The Causal Relation Hypothesis*

Concerning the causal hypothesis, if the relation established in Pavlovian conditioning is effectively a purely signal relationship, the assumption that in Pavlovian conditioning the animal learns a cause-effect relation is, at very the least, at odds with the idea that by means of learning the animal acquires a representation of its environment. Because to state that, in Pavlovian conditioning, causal relations are learned means that in this situation, animals learn causal relations that do not exist, but they do not establish merely signal relations despite the fact that these do occur. In addition to other conceptual difficulties involved, it seems clear that this assumption does not match the representative function of the environment that is attributed to conditioning.

At this point, it can be said that, for the causal hypothesis, the animals essentially go wrong with relations. They go wrong with signals because when they experience a purely signal relation -as in Pavlovian conditioning- they process it as though it were a causal relation. And they go wrong with causes because among the causes that, in the course of their experience, they consider as such, there will, in effect, be some causes but there will also be mere signals (such as those learned in Pavlovian conditioning according to the causal hypothesis).

Following this causal hypothesis the content of causal learning would be an unpalatable mixture in which effective causal relations, a reflection of a causal relation in the environment, will be confounded with spurious causal relations, that obey a mere spatial-temporal relation and are really signal relations. It is obvious that such a random and indistinct confusion of causes and non-causal signals in the animal's mind does not seem a very convincing way to depict the causal knowledge that allows the animal to represent its environment and to act consequently. (One could perhaps argue that the causal hypothesis only refers to Pavlovian conditioning and accepts that, in other situations of temporal contiguity, the animal would effectively process this relation and not a causal relation, as he is supposed to do in Pavlov's situation. Then, one would have to explain the differences between Pavlovian conditioning and other situations of temporal contiguity that bring about these differences in processing.)

Implicit in the above, although it deserves to be stated explicitly, we also find the effect -we might add, negative- that the causal hypothesis has on the meaning and value of the notion of signal. To propose that, in a situation of temporal contiguity, causal relations instead of signal relations are learned means that the signal relation is excluded from this situation and, therefore, from the operation of temporal contiguity. With no benefit in exchange, this exclusion puts the notion of signal on the spot.

Let us see: If in a situation of mere spatial-temporal contiguity of stimuli the animal does not establish a signal relation but a causal relation, then under what circumstances does it establish a signal relation? Not only do these circumstances remain unexplained, but it is very unlikely that they ever will be ascertained. Because, if the spatial-temporal covariation itself does not lead to the establishment of a signal relation, it is certainly hard to see what relation between stimuli -instead of or in addition to the covariation- could lead to an internal signal relation. Ultimately, given that no explanations are given about the way a signal is established or what environmental

relation it represents, one wonders whether the signal is assigned any place and function or whether it is a concept that can be discarded or considered beyond the associative explanations, which would obviously collide with a long and well-established tradition. Although, in that case, it would not matter very much whether or not animals go wrong with signals.

### *Hume's Influence*

This consideration of the causal relation hypothesis of Pavlovian conditioning would be incomplete if we did not take into account the close link between the causal hypothesis and the views of the British empiricist D. Hume (1739/1978). The causal hypothesis here discussed could be understood as application to Pavlovian conditioning of Hume's ideas about causal cognition. Readers are reminded that, according to Hume, in order to establish a causal relation between two events or stimuli, two external experiences or sensorial impressions are required: Spatial-temporal contiguity of the events and the precedence of one with respect to the other. But a third condition is necessary: The constant concurrence of the events, that is, the repeated experience of the contiguity and priority between the two stimuli.

According to Hume, this third -and main- condition of constant concurrence is not a new sensorial impression, from which the idea of causality could stem; it is rather an internal experience that, registered by the animal, has some effect on its mind. With constant concurrence, that is, by the objects being repeatedly experienced in contiguity and priority, the occurrence of the first object naturally brings to mind the idea of the second one. The mind takes note of its own propensity to pass from one object to the idea of another, and this observation or internal impression (that in the mind one object follows another) is represented in consciousness as a necessary connection or causal relation between the objects.

So, for Hume, the idea of necessary connection or causality that the mind attributes to the relation between objects is no more than the reflection or image in the consciousness of this observation or internal impression of the mind's activity. Necessity is just the effect of "that propensity, which custom produces, to pass from one object to the idea of its usual attendant" (Hume, 1739,1978, Bk.I, PtIII, Sect.XIV). So, one could say that, according to the British author, the sensory experience of the cause is reduced to the sensory experience of the signal: the sensations or sensory impressions that the notion of cause produces are the same as those that define the signal, to which the mind adds the internal observation of the way it operates. As indicated by Russell (1945), from Hume's doctrine follows that causes do not exist, despite which they are learned. This is tantamount to saying that causes have a mental life, despite not having a real life.

In any case, going along with Hume, it is entirely acceptable to say that in Pavlovian conditioning the animal learns or establishes a causal relation between stimuli. In effect, in the conditions of constant conjunction of a typical Pavlovian conditioning experiment, the mind passes from the conditioned stimulus to the image of the unconditioned stimulus and, when observing its own way of behaving -this passing

from one stimulus to the idea of another- the mind would get the idea of necessary connection between the stimuli and would make the first stimulus the cause of the second, independently of whether relation between the stimuli was one of purely spatial-temporal contiguity. That is the reason for the reluctance to enquire about the nature of the relation between the stimuli: In any case, the animal will learn a causal relation.

But the source of inspiration for the causal hypothesis is also the origin of the contradictions and difficulties that afflict it, as observed, and they derive essentially from the fact that Hume's explanation of how one learns a causal relation means that simple spatial-temporal contiguity is not processed. The mechanism or mental operation by which, for Hume, a stimulus is established as the cause of an effect necessarily implies that any stimulus that has a signal relation with another stimulus must be perceived not as a signal, but as a cause. There seem to be no exceptions to the rule that states that, given a stimulus that precedes another and evokes its image -that is, a stimulus that signals another stimulus- the mind will observe the internal sequence between the signal and the signaled and will reflect this internal experience as a necessity or causality. In other words, strictly according to Hume, any signal will start a mental operation that will lead to the inference of a causal relation, so that the representation of the signal disappears and reappears as the representation of something else. The cause is established at the cost of the pure signal, which as such seems, implicitly but effectively, to disappear or be eliminated from mental life.

Hume expressly says that animals learn causal relations (1739/1978, Bk.1, Pt.III, Sect.XVI): after experiencing the concurrence of stimuli, the animal mind also passes from one to the other and registers its own behavior. But this statement would be more precise if it said that animals establish causal relations but they cannot establish signal relations. According to Hume himself, when animals experience a spatial-temporal contiguity relation, ultimately, they do not process it as a contiguity relation, but as something else (as causality); given the appropriate circumstances to learn a signal, animals do not do so, establishing instead a cause. But if so happens that every signal ends up being processed as a cause, one can state that the signal is experienced but not learned -in contrast to the cause, which is not experienced but is learned- and thus, just as the cause was said to have a mental life but not a real one, one can say that, for Hume, the signal has a real life but not a mental one. Hence, in a situation of repeated contiguity of two events, animals do not experience causal relations, but they learn them (by carrying out some mental operation), and instead they experience signal relations, but they do not learn them (because they process them like causes). This diagnosis of Hume's is the antecedent and ultimate justification of the conclusion that, as seen, is deduced from the causal hypothesis and that states that, in Pavlovian conditioning, the animal establishes a causal relation between the conditioned stimulus and the unconditioned stimulus that does not exist, but it does not establish the signal relation that does exist. It is not necessary to emphasize that, in this way, neither Hume nor the causal hypothesis explain, first, how and under what conditions a causal relation that actually exists is processed, and second, how a signal that simply represents spatial-temporal contiguity is processed.

### ABOUT THE ADAPTIVE ARGUMENT

Lastly, and within the context of this reflection about the general issue of the conceptualization of conditioning in terms of signal or cause, it might be useful to refer again to the adaptive argument that accompanies the causal hypothesis and that, as mentioned above, assumes that it is advantageous for the animal's survival to detect and learn causal relations between events from its environment. There are doubtless adaptive benefits of causal learning, as there are also doubtless advantages of many other capacities, real or possible, but it is evident that the fact that a capacity is good does not mean that the animal has it or applies it in a given situation. Of course, it probably is adaptive for animals to learn or to be able to learn causes, but it is doubtless not adaptive for them to learn causes when they do not exist, as apparently -according to the causal hypothesis- would occur in Pavlov's situation.

conditioning also has -or, at least, so has it frequently been postulated- a far-reaching adaptive function. It seems clear that a simple signal allows the animal to represent an event from the environment in advance and thus prepare itself for it -either to obtain it or to avoid it- before it occurs. This same function can be attributed to the conditioned stimulus, about which, as a pure signal, it can consequently be said that it has an extensive "biological utility" (Hull, 1929; Pavlov, 1927). And indeed one could ask: What adaptive function has the causal relation that the establishment of signals has not? Because the function attributed to the causal relations acquired in Pavlovian conditioning seems to be no more than the prediction of the effect by the cause (Dickinson, 1980; Mackintosh, 1977). And the signal already has this function of anticipated representation or prediction of the signaled event, so that, ultimately, the function of causal knowledge and its alleged specific contribution to adaptation remains unclear.

Finally, one could argue that the establishment of pure signal relations between events from the environment may suffice to construct a significant representation that will contribute, to a great extent, to adaptation. Doubtless, causal knowledge comprises a precise and efficacious representation of the environment and opens up enormous action possibilities to its possessor. However, just as one could conceive a world of causes, one could conceive a world of spatial-temporal relations in which the knowledge of these relations, acquired by the animal when establishing signal relations, would be sufficient for the animal to make its way and survive. Perhaps, for the purpose of adaptation, the environment could be represented not as a "causal texture" but as a spatial-temporal net that the establishment of signals allows the animal to decipher.

### FINAL COMMENTS

Although perhaps somewhat obscured in the byways of its justification and development, the argument is essentially very simple. It makes no sense to say that, in Pavlovian conditioning, the animal learns causal relations because in Pavlovian conditioning, there is no such causal relation and therefore, the animal does not experience it. In Pavlovian conditioning, there is a purely signal relation between the stimuli, and

this is what one should assume that the animal learns. To accept the causal relation hypothesis leads to useless difficulties and contradictions that disappear then we regard the relation learned as a purely signal relation, based simply on spatial-temporal contiguity or covariance. This does not mean that animals cannot learn causal relations, but simply that they will not learn them when they do not exist and therefore, are not experienced.

No doubt, the conceptualization of conditioning as signal learning or as causal knowledge may not be very important. For example, neither one nor the other may be translatable to experimentally defined predictions, like Premack's (1993) statements about causal conceptualization. But in any case, as an explanation of Pavlovian conditioning, the notion of signaling seems preferable to that of causal knowledge, if only for the sake of coherence and a proper sense of concepts and terms.

### REFERENCES

- Dickinson A (1980). *Contemporary Animal Learning Theory*. Cambridge, England: Cambridge University Press
- Hull CL (1929). A functional interpretation of the conditioned reflex. *Psychological Review*, 36, 498-511.
- Hume D (1739/1978). *A Treatise on Human Nature*. Edited by LA Selby-Bigge & PH Nidditch, 2nd ed. Oxford: Oxford University Press.
- Mackintosh NJ (1977). Conditioning as the perception of causal relations. In RE Butts & J Hintikka (eds). *Foundational Problems in the Special Sciences* (pp. 241-250). Dordrecht: Reidel.
- Mackintosh NJ (1997). Has the wheel turned full circle? Fifty years of Learning Theory, 1946-1996. *The Quarterly Journal of Experimental Psychology*, 50A, 879-898.
- Morris CW (1946). *Signs, Language, and Behavior*. New York: Prentice-Hall.
- Pavlov IP (1927). *Conditioned Reflexes*. (GV Anrep, Trans.). Oxford, England: Oxford University Press. (Dover ed., 1960).
- Premack D (1993). Prolegomenon to evolution of cognition. In TA Poggio & DA Glaser (Eds.), *Exploring Brain Functions: Models in Neuroscience* (pp. 269-290). New York: Wiley
- Rescorla RA & Holland PC (1976). Some behavioral approaches to the study of learning. In MR Rosenzweig & EL Bennet (Eds.), *Neural Mechanisms of Learning and Memory* (pp. 165-192). Cambridge, Mass: MIT.
- Russell B (1945). *A History of Western Philosophy*. New York: Simon and Schuster.
- Testa TJ (1974). Causal relationships and the acquisition of avoidance responses. *Psychological Review*, 81, 491-505.
- Tolman EC (1932). *Purposive Behavior in Animals and Men*. New York: Irvington.

Received, December 26 2005  
Final acceptance, May 22 2006