

## Signalization and Stimulus-Substitution in Pavlov's Theory of Conditioning

Víctor García-Hoz  
Complutense University of Madrid

The concept of conditioning as signalization proposed by Ivan P. Pavlov (1927, 1928) is studied in relation to the theory of stimulus-substitution, which is also attributed to him. In the so-called theory of stimulus-substitution a distinction must be made between an empirical principle of substitution and an actual theory of substitution, which can adopt different forms. The Pavlovian theory of substitution—which conceives substitution as a substitution of the unconditioned stimulus (US) by the conditioned stimulus (CS) in the activation of the representation of the former—can be understood as an explanation or model of signalization. Signalization and substitution are answers to different questions, and the level of analysis to which signalization corresponds, is that which concerns the nature of conditioning as an operation of the animal in the environment.

*Keywords: Pavlovian conditioning, signalization, substitution*

El concepto del condicionamiento como señalización propuesto por Ivan P. Pavlov (1927, 1928) es estudiado en relación a la teoría de la sustitución del estímulo, también atribuida al fisiólogo ruso. En la así llamada teoría de la sustitución del estímulo debe hacerse una distinción entre un principio empírico de sustitución y una auténtica teoría de la sustitución, que puede adoptar formas diversas. La teoría pavloviana de la sustitución—que la concibe como una sustitución del estímulo incondicionado por el estímulo condicionado en la activación de la representación del primero— puede entenderse como una explicación o modelo de señalización. Señalización y sustitución son respuestas a diferentes preguntas y el nivel de análisis al que corresponde la señalización es el que concierne a la naturaleza del condicionamiento como una operación del animal en su medio ambiente.

*Palabras clave: condicionamiento pavloviano, señalización, sustitución*

I felt it appropriate to take advantage of the occasion we are celebrating—one of the first public presentations of Pavlov’s experimental and theoretical work—to try and deal with a “Pavlovian” question, one which may seem elemental and simple, but which is at the same time essential: the question of the description of conditioning. The specific question is: How does Pavlov conceptualize conditioning? Not so much in terms of the physiological mechanisms that support it or are responsible for it, but rather understanding it as a class of behavior or a significant activity of the animal in its environment. There are those that consider the question resolved, but for others it has been resolved unsatisfactorily. In any case, it seems appropriate, in paying tribute to a man of science, to concentrate on his theory here.

As a reader, I have always believed it to be evident in the writings of Pavlov that the concept of conditioning is inextricably linked to that of signalization and that, likewise, the concept of conditioned stimulus or reflex to that of signal or signaling stimulus or reflex. Reading *about* Pavlov more than reading Pavlov, however, I have also noticed that the concept of signalization is frequently neglected in descriptions of the Russian’s work. Most commonly, Pavlov’s name is linked to the concept—somewhat imprecisely—or the theory—somewhat ill-defined—of stimulus-substitution. And although in the dictionary signalization and substitution may have a shared meaning, it does not seem to be that which is denoted on speaking of “stimulus-substitution.” Rather, this expression would appear to be intended to denote something different, and even incompatible with signalization.

In short, I believe there is a need to clarify the relationship between the concepts of signalization and substitution in the explanation of conditioning—concepts, it must be said, that are perhaps contradictory and opposed, and that are nevertheless both embraced by Pavlov; and concepts whose analysis relates to the question of the original nature of conditioning. Before examining this relationship, I shall first concentrate on the evidence that, for Pavlov, conditioning is an operation of signalization of one stimulus by another. Subsequently, I shall try to clarify the meaning of the term “stimulus-substitution,” and in particular, what is meant by the theory of stimulus-substitution.

### Conditioning as Signalization in Pavlov

In the first and second lessons of *Conditioned Reflexes* (1927), his *opus magnum*, Pavlov, before dealing with the mechanics of conditioning, reflects upon its meaning as a form of action in animals. Pavlov aims to define conditioning in terms of an abstract class of behavior of the animal in the environment, as well as explaining its adaptive value. Perhaps what really interests him, above all, is the adaptive value of conditioning, but in order to deal with that aspect, Pavlov first attempts to define conditioning in terms of a class of abstract behavior or, in other words, as a type of

operational relationship between the animal and the environment.

In these first lessons, Pavlov conceptualizes conditioning as an operation of signalization of one stimulus by another, that is, an operation by means of which one stimulus is converted for the animal into a signal of the other. Literally, signalization is the mode of action of the brain by virtue of which the animal responds to one stimulus as a signal of the other. As Pavlov (1927) writes: “[W]e have found that a great number of all sorts of stimuli always act through the medium of the hemispheres as temporary and interchangeable signals for the comparatively small number of agencies of a general character which determine the inborn reflexes .... To this function of the hemispheres we gave the name of ‘signalization’” (pp. 16-17).

Signalization being a mode of action or response, conditioning is the operation or experience that makes possible that mode of action. If signalization is responding to one stimulus as the signal of another, conditioning is the establishment of a stimulus as the signal of another. Conditioning is an operation of the brain—which corresponds to an action of the environment—whose result is signalization as a mode of response. In other words, it is an experience processed by an animal which permits it to respond to a stimulus not so much for itself, for its particular and immediate properties, but rather because that stimulus is the signal for another.

Thus, the conditioned stimuli resulting from the operations of conditioning are, for Pavlov, stimuli whose essential property is that they signal the occurrence of another stimulus, and which are responded to in accordance with that other stimulus. Conditioned stimuli are “those stimuli which have the character of signals (*original italics*)” (Pavlov, 1927, p. 23). Pavlov insists on this character of conditioned stimuli:

[T]he normal animal must respond not only to stimuli which themselves bring immediate benefit or harm, but also to other physical and chemical agencies—waves of sound, light, and the like—which in themselves only *signal* the approach of these stimuli (*original italics*) .... Thus we see that the fundamental and the most general function of the hemispheres is that of reacting to signals presented by innumerable stimuli of interchangeable signification. (p. 14-15)

In general, in *Conditioned Reflexes*, Pavlov uses indistinctly the terms conditioned stimulus and signaling stimulus, as well as the terms conditioned reflex and signal reflex or signaling reflex.

The texts included in the anthology *Lectures on Conditioned Reflexes* (Pavlov, 1928) also show the constant importance this Russian master’s attributes to the idea of conditioning as signalization, as we can see in these quotations: a) “In this way numberless, various and remote external agents act as *signals* for food (*original italics*)”

(p. 123); b) “[T]his combination [of stimuli] comes to play the role of a *signal*, and becomes the conditioned stimulation” (p. 127); c) “[T]hese temporary stimuli become *signals*—the agents of the constant stimuli (original italics) (p. 214); d) [C]onditioned reflexes play the role of *signals* (original italics)” (p. 382).

And it is upon this signalization nature that, for Pavlov, the adaptive value of conditioning depends—the question of adaptive value being an essential one, and an issue that Pavlov considers necessary to deal with in order to justify the study of reflexes. For Pavlov (1927), conditioning is adaptive because conditioning is signalization. “[Signalization] is the only means by which a most delicate adjustment of the organism to the environment can be established” (p. 17).

In order to improve our understanding of Pavlov, and although the biological meaning of conditioning is not the object of this work, it is relevant to note that a concern regarding adaptation is present in Pavlov right from the start. The following is from the Madrid Lecture whose centenary we celebrate in this special issue (1903/1928):

[O]ne need only remind oneself of a case in which the saliva of certain species contains poison as a protection against other animals to see what a great significance for life this expectant production of the poison gland can have when enemies only approach, i.e., when they are first seen or heard (signaling action) .... By means of distant and accidental characteristics of objects the animal seeks his food, avoids enemies, etc. (p. 52)

Shortly afterwards, in the Huxley Lecture at London’s Charing Cross Hospital, Pavlov (1906/1928) stresses the fact that: “[I]n these conditioned stimuli, looked at from the point of view of general biology, we have a most perfect mechanism of adaptation .... The body has the capacity to react in a sensitive way to the phenomena of the outer world which are essential to it, because all other phenomena of the outer world, even the most insignificant, coinciding even temporarily with the essential become their indicators or, as they may be called, their signaling stimuli” (p. 87).

In sum, we find in Pavlov a theoretical principle of signalization that conceptualizes conditioning as a well-defined biological activity, that is, a given relational operation of the animal with the environment (signalization of stimuli) which fulfils a certain adaptive function. Upon this principle of signalization, Pavlov (1927) built the physiological study of conditioned reflexes: “Our next step will be to consider the question of the nature of signalization and of its mechanism from a purely physiological point of view” (p. 23). Having adopted this point of view, Pavlov attempts to show that signalization, the acquired response of an animal to a signal relationship, is a reflex response: “All these conditions leave no ground for regarding the phenomena which we have termed ‘signalization’ as being anything else than reflex” (p. 24).

## The Stimulus-Substitution Theory

The attribution to Pavlov of the stimulus-substitution theory (SST) is based on the fact that Pavlov repeatedly writes that the CR is equal or identical to the UR, and that, therefore, we can say that the CS “substitutes” the US, i.e., acts as a “substitute” for it in the production of the response (1927, 1928). If such affirmations constitute a theory, then it is clear that the SST must be attributed to Pavlov. However, if the SST is something more than the corroboration of a finding – real or assumed –, then it is doubtful whether the attribution is valid. On the one hand, the SST is sometimes understood as a simple set of empirical generalizations that Pavlov indeed proposes, to which are added a number of non-necessary consequences of these affirmations that nevertheless come to be represented as original conclusions of Pavlov and an effective part of his theory. On the other hand, it is certainly debatable whether Pavlov maintains some of the mechanisms and/or assumptions that are commonly known as the SST.

There is no canonical version of the SST—certainly not in the works of Pavlov—but there may be a standard version. It is noticeable that the SST has the peculiarity (albeit fairly unexceptional) of being a theory commonly attributed to others, and which is described in critical terms. The purpose of that which follows is to establish the terms of the most common or standard version of SST, examining the extent to which it can be attributed to Pavlov.

### *The Theoretical-Empirical Principle of Stimulus-Substitution*

One of the earliest attempts to identify a theory of stimulus-substitution was that of Hilgard (1936). For Hilgard, the ‘principle’, ‘concept’ or ‘formula’ of stimulus-substitution is identified with as “the fundamental concept that the conditioned stimulus acquires the potentiality of evoking the response originally elicited by the unconditioned stimulus” (p. 371). The same idea can be expressed by saying that, as the result of conditioning, the CS “substitutes” the US in the production of an identical and particular R (previously produced only by the US, and now produced also by the CS). Hilgard also writes that “in its most dogmatic expression, the formula of stimulus-substitution demands that an anticipatory conditioned response and a following unconditioned response are successive reactions of essentially the same kind” (p. 381). Hilgard does not, however, clarify the differences between the dogmatic expression and other seemingly possible expressions.

More modern descriptions of the SST do not contradict Hilgard’s. Hearst and Jenkins (1974) write: “The orthodox, Pavlovian theory of classical conditioning is *stimulus substitution*. The theory holds that as the result of pairings, the CS comes to evoke a response greatly resembling that

evoked by the US. The general observation on which the theory rests is the close similarity between the conditioned response and the unconditioned response" (original italics) (p. 33). For Le Ny (1983), the "principle of substitution" is the "explanatory principle of classical conditioning according to which the conditional stimulus, after its conjunctions with the unconditional stimulus, takes the place of the latter in its effect on the reaction" (p. 239). Le Ny expressly warns that the principle of substitution is "a theoretical notion, not a purely operational one" (p. 239).

Mackintosh has explored the theory of stimulus-substitution at length, and his analyses have become a *locus classicus* in the study of the matter. Mackintosh (1974), in explaining the theory of stimulus-substitution, writes "Pavlov's account of what is happening here [in a classical conditioning experiment] was that an initially indifferent stimulus, the CS, comes to elicit the response belonging to another stimulus, the UCS: one stimulus becomes a substitute for another" (p. 79). Elsewhere, Mackintosh (1983) writes: "Conditioning, for Pavlov, was simply a matter of the metronome coming to substitute for the food, and thus itself eliciting salivation" (pp. 4-5).

A certain confusion within the levels of interpretation can be discerned in the above. On the one hand, the SST is identified with the empirical fact—real or supposed—that the response to the conditioned stimulus is identical or quite similar to the response to the unconditioned stimulus. From this point of view, and despite its name, the SST seems more like a generalization or empirical principle than an actual explanation or theory. On the other hand, however, and without adding anything to this same empirical fact, certain authors refer specifically to the theoretical nature of the principle of substitution, as though from this fact there arose by itself a particular form of explaining conditioning, which would be the SST.

It is true that the fact of substitution must be explained, in other words, the CR-UR similarity needs a theory of conditioning that accounts for it. From this perspective, it would be correct to state that the SST is any theory that attempts to explain conditioning on the basis of the assumption that the CR is always, or by its nature, the same as the UR. But it is not in this sense that one commonly speaks of the "stimulus-substitution theory;" rather, what is generally understood by this is what should more accurately be called the "pre-stimulus-substitution theory." Not so much a theory that attempts to explain conditioning on the basis of substitution, but rather a judgment about what an SST should be like. There is no doubt that the CR-UR similarity is (or would be) a fact that requires a theory and restricts it, but does not take it wholly for granted, and leaves room for various possible explanations. In accordance, however, with the pre-stimulus-substitution theory, accepting the empirical CR-UR similarity necessarily involves a particular form of explaining that similarity, and, in the end, conditioning itself.

Thus, Hilgard (1936) writes that the principle of stimulus-substitution is "the orthodox expression of conditioning," as opposed to an interpretation that is "organismic, sign-learning, or Gestalt" (p. 370), which Hilgard considers preferable. One of the main differences between the two points of view resides in the fact that the former is "in essence mechanistic," while the latter is "essentially teleological" (p. 370). Fifty years later, Rescorla (1988) would write that: "describing Pavlovian conditioning as the endowing of a CS with the ability to evoke the same response as the US is a wholly inadequate characterization" (p. 158). This inadequate description is attributed to "the reflex tradition in which Pavlov worked and within which many early behaviorists thought" (p. 152), stressing that this tradition "sees conditioning as a kind of low-level mechanical process in which the control over a response is passed from one stimulus to another" (p. 152). Rescorla contrasts this interpretation with that which, following Tolman, considers conditioning as a cognitive operation that "involves the learning of relations among events" (p. 158).

Miller and Matute (1996), somewhat more eclectic, suggest that in Pavlovian conditioning it may be necessary to distinguish between a "mechanistic" form of acquisition and a form of acquisition in which "the critical learning can be regarded as a purely predictive relation ... such as a cause has to an effect" (p. 134). "Mechanistic" acquisition would permit an explanation in terms of stimulus-substitution, which nevertheless would not be applicable to the acquisition of a predictive relationship.

In particular we find that, for historical and conceptual reasons, the SST is not uncommonly identified with stimulus-response (S-R) theory. That is, it is assumed that the only possible explanation of the empirical fact of substitution is to be found within the framework of the S-R theory. In accordance with the S-R theory, it should be remembered, learning occurs because a stimulus that acts on an animal at the moment at which the animal gives a response is associated with that response and is henceforth capable of producing it; in conditioning, the CS is present when the original response (UR) to the US occurs, so that a stimulus-response association is established between the CS and the UR by virtue of which the former now produces the latter.

Regardless of circumstantial or historical reasons (see Boakes, this issue), there are intrinsic reasons that favor the identification of the SST with S-R theory. On the one hand, it is clear that the SST coincides with S-R theories in the claim for CR-UR similarity, and the differences between the CR and the UR, which constitute a problem for the SST, are also problematic for S-R theories (Mackintosh, 1974, 1979). Furthermore, the S-R theory or mechanism just described—and which corresponds, for example, to Guthrie's notion that the responses which are conditioned are the responses which occur—explains without much difficulty the empirical fact of substitution: The CS produces the

particular activity that has occurred in its presence and with which it has been associated, and which is none other than the original response to the US.

Thus, we can speak of the SST in terms of any of the meanings that may be found in an empirical and theoretical principle which, in its narrowest sense, is restricted to affirming the similarity between the CR and the UR, and which, in its widest sense, takes for granted that an explanation of conditioning necessarily results from that similarity. The distinction between these two views is not always made, and when it is, it is not always clear and precise.

Understood as an empirical principle, the SST is legitimately attributed to Pavlov, who, as indicated above, accepts as a fact, and attempts to explain, that the CR is the same as the UR. What is more doubtful is whether we can attribute certain theoretical concepts to Pavlov that tradition and confusion link to the acceptance of the empirical fact of substitution. It is difficult to agree, for example, with Mackintosh's (1983) claim that, for Pavlov, "the essence of conditioning" is substitution (p. 5). Similar claims are more than frequent. What is important to realize is that, if S-R theory accounts for CR-UR similarity, this similarity can also be explained from the point of view which the classical division between theories of learning considers as the opposite, that of the stimulus-stimulus theory. And this is, ultimately, the explanation or theory preferred by Pavlov.

### *Theories of Substitution*

The explanation of substitution as an empirical fact can be dealt with more appropriately by distinguishing between theory of learning and theory of performance. The theory of learning attempts to respond to the question of the "content" or "elements" of learning; it can be said that, if conditioning is an experience, the theory of learning tries to explain what that experience consists of, or what its components are, and how it is registered. The theory of performance, on the other hand, refers to the way the learning experience manifests itself in action, the way the learned response is produced. As far as both learning and performance are concerned, a classic division is established, as referred to above, between S-R theories and S-S theories.

According to the S-R theories of learning, as mentioned above, conditioning occurs because the CS is present upon the production of the UR; in other words, CS and UR are the main elements of the experience, and between them there is established a stimulus-response association. According to the S-S theories of learning, on the other hand, the significant elements of the conditioning experience are the CS and the US; learning occurs because a first stimulus that acts upon the animal at the moment at which it is experiencing another stimulus is associated with this second stimulus—that is, it is associated with the activation of the representation or center of this second stimulus. In conditioning, the CS is present when the US occurs, so that

between the CS and the US an association stimulus-stimulus is established, which means that the presentation of the CS will henceforth elicit the activity of the center or representation of the US in its absence. Pavlov's theory of learning (Pavlov, 1927) is an E-E theory, according to which the CS connects with the "cortical center" or nervous representation of the US.

It can be proposed that the theory of learning has nothing directly to do with the fact of substitution, which should be explained—insofar as it needs to be—within the framework of the theory of performance. Mackintosh (1974) wrote that "this 'stimulus-substitution' theory ... does not, however, say anything about the nature of the associations formed" (p. 79), which is equivalent to saying, or can be understood to mean, that the SST does not exist as a theory of learning.

However, in no way is the theory of performance independent from the theory of learning; rather, the hypotheses on the generation of the response, in which substitution should perhaps be explained, depend on the way in which the learning experience is described. Thus, corresponding to the S-S and S-R theories of learning, are isomorphic theories of performance that directly confront the fact of CR-UR similarity. Mackintosh's argument should be understood not in the sense that the theory of learning does not deal with stimulus-substitution, but rather in the sense that it offers various versions of this substitution. It could be said that, for the S-R theory of learning, substitution is attributable to the fact that the CS connects with the response to the US, while the S-S theory refers to the fact that the CS connects with the activation of the center or representation of the US.

As far as the theory of performance is concerned, it has been stated that the assumption that the CS acquires the capacity to elicit the response originally elicited by the US implies "a highly restricted view of how conditioning affects behavior," and allows the concept of "only one way in which performance is generated" (Rescorla, 1988, pp. 156-157). However, as pointed out above, there are at least S-S and S-R theories on the performance of the learned response, isomorphic with the mentioned theories of learning, which explain substitution in a different way.

For the S-R theory of performance, the similarity between the CR and the UR arises naturally from learning (as the S-R theory itself understands it): the CS produces the R that was originally produced by the US because the occurrence of the CS excites directly the R that previously occurred in its presence (and with which it has become associated or connected in the learning process). The CR is the UR itself, now produced by another stimulus.

In the S-S theory of performance proposed by Pavlov, the CS produces the response that was originally produced by the US because the occurrence of the CS activates the center of the US with which it has become associated or connected during the learning process, and on activating this center of the US, it also excites the responses that the occurrence of the US normally produces. Mackintosh (1983)

uses the term “stimulus-substitution theory” to refer to the theory of performance in conditioning, first proposed by Pavlov, which is formulated by stating that, after learning: “presentation of the former [the CS] will activate a representation of the latter [the reinforcer]; but activating the representation of the reinforcer by presenting the reinforcer itself normally elicits a particular set of responses; it follows that the new, indirect way of activating this representation of the reinforcer should also elicit this set of responses. The CS should therefore elicit responses similar to those elicited by the US” (p. 51).

There are, then, two stimulus-substitution theories, that is, two theories which, taking stimulus-substitution as an empirical fact that requires an explanation, find this explanation—both at the level of learning and at that of performance—in the operation of hypothetical mechanisms that are different and in some ways contradictory. The S-R theory and the S-S theory may commit the same initial error, and in some cases make the same mistaken predictions, but they explain substitution and, ultimately, conditioning, in different and even opposing ways, so that, while there are criticisms that affect them both, not all the criticisms that are made of one can be to the other; nor are the limitations of one necessarily the limitations of the other.

Upon explaining substitution by means of different mechanisms, each one of these theories lends a different sense to the claim that the CS substitutes the US. According to an S-R theory of substitution, the CS substitutes the US in the sense that, in learning, the CS connects with the response that the US originally produces. In terms of performance, the CS substitutes the US in the sense that the CS excites, in the absence of the US, the original response to the US. In contrast, according to Pavlov’s S-S theory of substitution, the CS substitutes the US insofar as, in learning, the CS connects with the excitation of a center originally produced by the US; and in terms of performance, the CS substitutes the US in the sense that the CS elicits, in the absence of the US, the center or representation originally elicited by the US.

It can be seen that, contrary to the S-R theories, the S-S theory of substitution assigns a critical function to the US, in both the formation and registration of the association and the production of the response. The US operates in the establishment of the association, which takes place because the US occurs after the CS, and the activity of its center or representation is one of the elements that remains associated with, and determines the functioning of, the association and the production of the response.

### Signalization and Substitution

Having identified a principle of signalization and a principle of substitution, which it could be said were maintained by Pavlov, it is now necessary to examine the relationship—perhaps one of contradiction or opposition—

between the two. The aim is, on the one hand, to determine if it makes sense to maintain that conditioning is an operation of signalization of the US by the CS, characterizing the response to the CS as a response to its status as a signal of the US, and, on the other hand, to show that the response to the CS is the same as the response to the US, so that the former acts as a substitute of the latter.

The relationship between signalization and substitution is not an issue that has been dealt with by many authors, mainly because few authors have dealt with the concept of signalization, which is not even commonly referred to (though similar concepts are found) among the critics of substitution. A notable exception is that of Tolman (1932), who provided inspiration for the present work.

In any case, it can be argued that there is a point of view according to which signalization and substitution would be opposed to one another as conceptions of conditioning. The question is the nature of the mechanism to be considered in each case. Thus, substitution, the fact that a stimulus present when a response is produced is henceforth capable of producing it, seems easy to understand and represent in mechanical terms, while from the cognitive point of view it is somewhat impenetrable. Signalization, on the other hand, seems difficult to describe in mechanical terms, yet can easily be conceptualized as a genuine operation of knowledge. It is easy to argue that registering a stimulus as the signal of another and responding to a stimulus in this capacity are cognitive activities in the sense that they depend on the perception of a relationship and therefore require knowledge of that relationship. Thus, substitution would permit a mechanical interpretation of Pavlovian conditioning, which many authors consider, as well as insufficient, contradictory to the conceptualization of conditioning as a cognitive activity, in which, we might add, signalization would easily be accommodated. It seems to be understood, in sum, that a low-level mechanical process, perhaps acceptable as an explanation of the transfer of a reflex response, cannot explain learning in the context of the relationships between environmental facts.

Nevertheless, the relationship between signalization and substitution can be examined from more precise points of view.

### *The Principle of Signalization and the Empirical Principle of Substitution*

It is interesting to compare the principle of substitution with the principle of signalization in terms of the basic claim that the CR is identical to the UR. We should bear in mind, however, that it is not possible to make a direct comparison in these terms. For if substitution is an empirical principle, or, let us say, an empirical and theoretical principle, in any way linked to the fact of the similarity between the CR and the UR, signalization, in contrast, is a purely theoretical

principle that describes the operation of conditioning in certain terms, but does not involve precise predictions with respect to the relationship of similarity between the CR and the UR. Thus, in order to make this comparison it would seem necessary, first of all, to resolve the more general question of how a signal is responded to, and, specifically, the question of the similar relationship between the response to the signal and the response to that which is signaled.

It is obvious that there would be no contradiction between signalization and substitution, rather the contrary, there would be implication or consequence, based on the assumption that a signal is responded to in the same way as is the thing signaled. Pavlov perhaps commits the error of assuming that a signal is responded to in the same way as the thing signaled. However, there would be a contradiction if we were to consider—because of the very nature of the signal, which implies that it is something different from what it signals, or for any other reason—that the reaction to a signal should be different from the reaction to that which is signaled. On the other hand, if the response to a signal were sometimes the same as the response to that which is signaled and sometimes different, the principle of stimulus-substitution would be not so much erroneous as incomplete, on describing as a general case one that is particular (that the CR is the same as the UR), but one that is all the same possible from the point of view of signalization. In sum, there would only be incompatibility between one principle and another if it was assumed that a signal cannot be responded to in the same way as the thing itself.

There is possibly an intuitive inclination to assume that the response to a signal is sometimes the same as the response to that which is signaled and sometimes different, depending on the particular stimulus that acts as a signal and the particular stimulus that signals. Something along these lines appeared to be in Tolman's (1932) mind when he wrote: "And in this case, it so happened that the response originally made to the significate was also appropriate to the sign, as a temporal signal of this coming significate" (p. 331). Tolman clarifies by adding, "if it is physiologically good to secrete saliva for actual food, it is also good to secrete it just before the approach of food."

The idea that, depending on a series of circumstances, the response to a signal may be the same as or different from the response to that which is signaled can be linked to what seems to be the empirical fact that there are situations in Pavlovian conditioning in which the CR is similar to the UR and there are situations in which it is different. Thus, first of all, the Pavlovian interpretation of conditioning as signalization is strengthened. And second, it becomes clear that the empirical principle of substitution, being, as we said, erroneous in that it is too restrictive, is not incompatible—rather the opposite—with the principle of signalization, of which it can be said to constitute a possibility. Critics of substitution insist on the need to explain that the CR is not always, nor necessarily, the same as the

UR; such an explanation is undoubtedly necessary, but it is equally necessary to explain that the CR is sometimes the same as the UR.

A theory of signalization should, then, concern itself with substitution and explain it, correcting it or restricting it, perhaps, but in no way denying it. This is what Tolman (1932) does; his theory of Pavlovian conditioning as the learning of signs easily accommodates substitution: "So that the acquisition of the sign-gestalt-expectation in this case demonstrates itself in the carrying over of one and the same response. That is, the response made after learning to the total sign-gestalt-expectation is the same as the response made before learning to the significate-object by itself" (p. 331). There is certainly no contradiction in stating that, in conditioning, the CS is converted into a US signal, so that the form of response to the CS is determined by the US, with the CR and UR being identical in some circumstances and not in others. Signalization is the first principle in the explanation of conditioning and antecedes, so to speak, substitution, which should be understood as a subordinate principle.

### *Signalization and Theories of Substitution*

The relationship between signalization and substitution can also be analyzed from the point of view of the mechanism, and with the purpose of determining whether a theory that explains the latter can also be valid for the former.

From the point of view of the theory of learning, that is, the identification of the experience in which conditioning consists, the principle of signalization seems difficult to explain via an S-R theory, which can be considered as limited when compared to substitution. In fact, according to the principle of signalization, it is possible for the CR to be different from the UR, while the absence of similarity between the CR and the UR constitutes a serious problem for the S-R theory.

In contrast, the S-S theory of learning is the natural context for the principle of signalization, which it easily fits into. Pavlov's theory whereby the excitation of one center follows that of another, so that the first connects with the second, which explained substitution, also constitutes a way—and an irreproachable and almost inevitable way—of explaining signalization. At the level of stimuli and centers of sensory and motor activity, characteristic of Pavlov and of a large part of the theories of conditioning, there does not seem to be any other way of representing how one stimulus is linked to another—this link of one fact with another fact being what endows the former with the character of a signal.

It is also the case that Pavlov's theory of performance (the CR is produced because the CS excites the center of the US), which permitted the explanation of substitution, also—and perhaps within a single concept—explains signalization. The participation of the activity of the center

of the US in the CR constitutes an optimum way of representing that the CR depends on the US. That is, from the point of view of signalization, that the response to the signal depends on the thing signaled. The same can certainly not be said of an S-R theory of performance, which may explain substitution but would be hard pressed to explain signalization, since it would be difficult to maintain that the CR is the response to a signal when the US (or its center or representation) in no way participates or intervenes in the CR. The problem for the S-S theory may reside in its assumption that a CR and a UR, which can be different, arise, nevertheless, from the excitation of the center of one and the same US. The problem can be resolved by postulating different modes of activation of the center of the US, to which different forms of responding would correspond. This is what is proposed by modern theories of performance, which, acknowledging the absence of necessary similarity between the CR and the UR, do not stray from the Pavlovian explanation of performance, and attribute the CR to the activation by the CS of the representation of the US and, consequently, of the response processes dependent upon this response (Wagner & Brandon, 1989).

In sum, Pavlov's theories, proposed on the basis of substitution, and which in any case explain it, also constitute a sufficient explanation of signalization as an operation by which one stimulus is converted into a signal of another, and by which one stimulus is responded to as a signal of another. The signal's character of reference to another fact is taken into account, at the level in question, on assuming that learning, first, depends on the activity of the center of the CS (the signal) being linked to the activity of the center of the US (that which is signaled); and that performance, subsequently, depends on the activity of the center of the CS producing activity in the center of the US. This is a mechanism which results in the US substituting the CS, not the CS substituting the US, just as it can be said that what is signaled substitutes the signal in the sense that the latter is responded to in accordance with the former.

### Final Remarks

Given that, for Pavlov, conditioning was first and foremost an activity of the animal through which it takes one stimulus as the signal of another, it is interesting to note the customary absence of the concept of signalization in references to his work. Rescorla (1988) is right when he points out that the characterization of conditioning found in many textbooks is inadequate (Rescorla cites some general psychology and psychopathology texts, but the same could be said of the most well known texts on conditioning and learning). But Rescorla is wrong in linking this description to Pavlov, because much of the inadequacy of the textbooks is due to the omission of the essentially Pavlovian concept that is signalization.

The most common term in relation to Pavlov is that of stimulus-substitution, which Pavlov indeed also employs. Naturally, the notion of substitution can be understood in different senses. Conceived as substitution of the US by the CS in the excitation of the center of the US, which is the sense in which Pavlov's approach can be interpreted, stimulus-substitution and stimulus-signalization are not contradictory concepts, but should rather be understood as concepts that correspond to different levels of analysis. Stimulus-signalization as an operation of the animal is the first or main theoretical principle, while substitution can be understood as a derivative principle, or as a kind of explanation of signalization. In the study of conditioning, signalization antecedes, so to speak, substitution, so that to explain it we must begin by saying that it is an operation of signalization and, from then on, deal with substitution, as a subordinate principle.

This may become clearer in the context of the levels of analysis model proposed by Marr (1982). In Marr's terms we could say that signalization is a response to the question of "what" an organism or system does, as opposed to the question of "how" it does it. We should bear in mind that Marr proposes a distinction between the level of the operation or task the organism carries out and the levels related to the way in which that operation is performed. Marr writes that neglecting this distinction between levels, leading to confusion in these questions, has affected the progress of artificial intelligence; the same could be said for the study of conditioning.

Understanding the operation of the organism as computation, Marr uses the term "computational level" to refer to the first question or level, which involves the analysis of the computation the system performs, as well as the function of this computation. It would seem that it is to this computational level that Pavlov's characterization of conditioning—as a phenomenon of signalization of stimuli to which a biological function can be attributed—corresponds. Substitution, in the sense described, would be a concept belonging to the second level, that is, the level of how the organism does what it does. Signalization, it can be said, is obtained by means of a mechanism of substitution of the US by the CS in the activation of the center or representation of the former. In this context, it is probably important to note that signalization is understandable and makes sense as an operation with a certain function, which cannot be said of substitution. Substitution is an operation whose sense is not clear. In order to comprehend it, it is necessary to seek and attribute to it a content and a function, a function which may eventually be found in signalization.

An important feature of Pavlov's characterization of conditioning resides in the very fact that the question of the nature of conditioning is considered in terms of an operation of the animal vis-à-vis its environment. In order to understand conditioning it is necessary to characterize it in such terms;



otherwise, we would be ignoring an essential reality of the phenomenon and a level of analysis that must be taken into account. Perhaps signalization is not the answer, or there is a more sophisticated answer, but in any case there will always remain the question about conditioning as an operation that the Russian master tried to answer. In one way or another, Pavlov's interpretation of conditioning cannot be understood without the concept of signalization; and ultimately, we cannot understand conditioning without discussing this concept.

### References

- Boakes, R. A. (this issue). *The impact of Pavlov on the psychology of learning in English-Speaking countries.*
- Hearst, E., & Jenkins, H. M. (1974). *Sign-tracking: The stimulus-reinforcer relation and directed action.* Austin, TX: The Psychonomic Society.
- Hilgard, E. R. (1936). The nature of the conditioned response: I. The case for and against stimulus-substitution. *Psychological Review*, 43, 366-385.
- Le Ny, J. F. (1983). *El condicionamiento y el aprendizaje.* Barcelona, Spain: Península.
- Mackintosh, N. J. (1974). *The psychology of animal learning.* London: Academic Press.
- Mackintosh, N. J. (1983). *Conditioning and associative learning.* Oxford, UK: Clarendon.
- Marr, D. (1982). *Vision.* San Francisco, CA: Freeman.
- Miller, R. R., & Matute, H. (1996). Animal analogues of causal judgment. In D. R. Shanks, K. J. Holyoak, & D. L. Medin (Eds.), *Causal learning. The psychology of learning and motivation* (Vol. 34, pp. 133-166). San Diego, CA: Academic Press.
- Pavlov, I. P. (1927). *Conditioned reflexes.* Oxford, UK: Oxford University.
- Pavlov, I. P. (1903/1928). Experimental psychology and psychopathology in animals. In I.P. Pavlov (Aut.), *Lectures on conditioned reflexes* (Vol. I, pp. 47-60). New York: International.
- Pavlov, I. P. (1906/1928). Scientific study of the so-called psychical processes in the higher animals. In I.P. Pavlov (Aut.), *Lectures on conditioned reflexes* (Vol. I, pp. 81-96). New York: International.
- Pavlov, I. P. (1928). *Lectures on conditioned reflexes* (Vol. 1). New York: International.
- Rescorla, R. A. (1988). Pavlovian conditioning: It's not what you think it is. *American Psychologist*, 43, 151-160.
- Tolman, E. C. (1932). *Purposive behavior in animals and men.* New York: Appleton-Century-Crofts.
- Wagner, A. R., & Brandon, S. E. (1989). Evolution of a structured connectionist model of Pavlovian conditioning (AESOP). In S. B. Klein & R. R. Mowrer (Eds.), *Contemporary learning theories: Vol. 2. Pavlovian conditioning and the status of traditional learning theory* (pp. 149-189). Hillsdale, NJ: Erlbaum.

Received June 11, 2003

Revision received September 30, 2003

Accepted October 2, 2003