

# A response rule for positive and negative stimulus interaction in associative learning and performance

OSKAR PINEÑO

*Hofstra University, Hempstead, New York*

The present article presents a response rule developed to account for both positive and negative stimulus interaction. In the response rule proposed here, positive interaction phenomena (e.g., second-order conditioning) and negative interaction phenomena (e.g., Pavlovian conditioned inhibition) are presumed to occur during performance and acquisition, respectively. Also, in this rule the novelty of the test stimulus determines the expression of positive interaction on responding. As the stimulus loses its novelty over training, positive interaction effects will wane, which will allow negative interaction effects to emerge in responding elicited by the stimulus. It is proposed that this response rule can be adopted by acquisition-focused associative models (e.g., Rescorla & Wagner, 1972) in order to account for contrary associative phenomena in the literature. The simulation program used in this study is available for download from the author's website and from the Psychonomic Society website at [www.psychonomic.org/archive](http://www.psychonomic.org/archive).

The conditioned response elicited by a target cue or conditioned stimulus (CS) due to its pairings with an outcome or unconditioned stimulus (US) is known to depend on the associative status of other CSs that were presented in compound with the target CS. In fact, training two CSs in compound and observing the impact of the associative history of the nontarget CS on responding to the target CS (i.e., what is known as *stimulus interaction*) is perhaps one of the most powerful tools for the study of the processes involved in associative learning. Generally speaking, two main types of stimulus interaction can be observed after compound treatment, which we shall refer to as positive and negative interaction. In *positive interaction* phenomena the response potential of the target CS is positively related to the response potential of the nontarget CS. This is the case of second-order conditioning (Pavlov, 1927), sensory preconditioning (Brodgen, 1939), and representation-mediated acquisition and extinction (e.g., Holland, 1981; Holland & Forbes, 1982). In *negative interaction* phenomena the response potential of the target CS is negatively related to the response potential of the nontarget CS. Examples of negative interaction are Pavlovian conditioned inhibition and stimulus competition effects, such as overshadowing (Pavlov, 1927) and blocking (Kamin, 1968).

With some notable exceptions (e.g., Holland, 1981), most models of learning have ignored the need for explanations of positive interaction phenomena and have exclusively focused on accounting for negative interaction effects. This is the case of associative models of learning (e.g., Dickinson & Burke, 1996; Mackintosh, 1975; Pearce & Hall, 1980; Rescorla & Wagner, 1972; Van Hamme &

Wasserman, 1994; Wagner, 1981), statistical models (e.g., Allan, 1980; Cheng, 1997), and comparator theory (Denniston, Savastano, & Miller, 2001; Miller & Matzel, 1988). This inability of models of learning to account for positive interaction becomes a more serious problem when one considers that highly similar experimental treatments can yield either positive or negative interaction effects. One such treatment consists of presenting pairings of the nontarget CS, A, with the US (i.e., A→US trials) interspersed with nonreinforced presentations of A in compound with the target CS, X (i.e., AX→noUS trials). This treatment can result in CS X eliciting either excitatory responding (i.e., second-order conditioning) or behavior indicative of inhibition (i.e., Pavlovian conditioned inhibition). Interestingly, some studies found that second-order conditioning occurs after few AX→noUS trials, followed by the development of Pavlovian conditioned inhibition after many AX→noUS trials (Stout, Escobar, & Miller, 2004; Yin, Barnet, & Miller, 1994). These authors (also see Gewirtz & Davis, 2000) interpreted these results as indicative of the operation of two processes that develop at different rates. Specifically second-order conditioning presumably develops faster than Pavlovian conditioned inhibition and, as a consequence, excitatory responding to X prevails after few AX→noUS trials. However, after many nonreinforced presentations of the AX compound, CS X accrues enough inhibitory strength to counteract the expression of second-order conditioning, allowing inhibitory responding to prevail after many AX→noUS trials.<sup>1</sup>

The observation of a transition from second-order conditioning (positive interaction) to Pavlovian conditioned

---

O. Pineño, [oskar.pineño@hofstra.edu](mailto:oskar.pineño@hofstra.edu)

inhibition (negative interaction) by Yin et al. (1994; see also Stout et al., 2004) recently encouraged Stout and Miller (2007) to develop a mathematical formulation of the extended comparator hypothesis (Denniston et al., 2001) that can account for both positive and negative interaction phenomena. As in the original and extended comparator hypothesis, in Stout and Miller's SOCR (i.e., Sometimes Competing Retrieval) model, the presentation of the target CS at test will retrieve the memory of the US-representation both directly and indirectly, through the activation of the representation of those CSs previously trained in compound with the target CS. However, contrary to what the original and extended comparator hypothesis assumes, in SOCR the indirect activation of the US-representation does not necessarily impair responding to the target CS. Rather, the indirect activation of the US-representation can sometimes facilitate responding to the target CS. A variable, the "switching operator," determines the kind of impact (i.e., facilitative or competitive) the indirect US-representation will have on responding produced by the target CS. Specifically, after few training trials the switching operator promotes stimulus facilitation (positive interaction) and, after many training trials, stimulus competition (negative interaction).

This article introduces a response rule similar to the one proposed by Stout and Miller (2007), which also aims to explain the transition from positive to negative interaction that is usually found as a function of the number of training trials with the target CS. The critical difference between the present rule and that of the SOCR model is that the response rule proposed here is compatible with any associative model (e.g., Mackintosh, 1975; Pearce & Hall, 1980; Rescorla & Wagner, 1972; Wagner, 1981), whereas the SOCR model departed from the extended comparator hypothesis and therefore depends on the assumptions of such model. Thus, the present article does not introduce a new model, but a response rule that could be incorporated in many existing associative models, allowing them to account for a wider range of phenomena. In fact, the ultimate goal of this article is to show that the explanation of the transition from positive to negative interaction does not require the assumptions of the extended comparator hypothesis, as implemented in the SOCR model: Any traditional associative model, such as the Rescorla–Wagner (1972) model, could account for this effect by adopting a response rule like the one proposed here.

### The Rule

In the present response rule, the response elicited by a CS,  $X$ , is assumed to depend not only on the associative strength of  $X$ , but also on the associative strength of other CSs,  $i$ , that are associatively activated by  $X$ . This is represented in Equation 1:

$$R_X = V_{X-US} + N_X \cdot \sum V_{X-i} \cdot V_{i-US}. \quad (1)$$

In this equation,  $R_X$  represents the conditioned response elicited by CS  $X$ ,  $V_{X-US}$  denotes the associative strength of CS  $X$  (i.e., the strength of the  $X$ –US association),  $V_{X-i}$  denotes the strength of the within-compound association between  $X$  and each of  $X$ 's companion CSs,  $i$ , and  $V_{i-US}$

denotes the associative strength of each of  $X$ 's companion CSs (i.e., the strength of the  $i$ –US associations). A variable of critical importance in the response rule of Equation 1 is  $N_X$ , which represents the perceived novelty of CS  $X$ . The values of  $N_X$  range from 1 to 0 (i.e.,  $N_X = 1$  when  $X$  is presented for the first time, and  $N_X = 0$  when  $X$  is completely familiar). The value of  $N_X$  will decrease toward 0 on each trial  $X$  is presented, according to Equation 2A:

$$\Delta N_X^t = \alpha_X \cdot n_{\text{rate}} \cdot (0 - N_X^{t-1}). \quad (2A)$$

According to this equation, the value of  $N_X$  will decrease on trial  $t$  as a function of the difference between 0 and the value of  $N_X$  on the previous trial ( $t-1$ ), and proportionally to the salience of  $X$  ( $\alpha_X$ ) and a novelty-rate parameter ( $n_{\text{rate}}$ ). Both  $\alpha_X$  and  $n_{\text{rate}}$  adopt values between 0 and 1. The value of  $\alpha_X$  depends on the physical intensity of  $X$  (i.e., as in the Rescorla–Wagner model), whereas the value of  $n_{\text{rate}}$  depends on certain aspects of the experimental situation, such as the experimental preparation and/or animal species (for the sake of simplicity, the value of  $n_{\text{rate}}$  is assumed to be identical for all different CSs). Finally, the decrement of  $N_X$  on trial  $t$  is added to the value of  $N_X$  on trial  $t-1$ , as shown by Equation 2B.

$$N_X^t = N_X^{t-1} + \Delta N_X^t \quad (2B)$$

### Meaning of the Rule

The response rule represented in Equation 1 simply states that responding to a CS,  $X$ , can depend not only on the associative strength of  $X$ , but also on the associative strength of those CSs,  $i$ , that are associatively activated by the presentation of  $X$ . In this rule, the expression of the associative strength of CSs  $i$  is modulated by the novelty of  $X$ . In order to better understand the present response rule, some of its features need to be explained.

First, the acquisition of associative strength is assumed to be a competitive process and, thus, the response rule can incorporate any of the learning rules proposed by acquisition-focused associative models (e.g., Dickinson & Burke, 1996; Mackintosh, 1975; Pearce & Hall, 1980; Rescorla & Wagner, 1972; Van Hamme & Wasserman, 1994; Wagner, 1981). That is, CSs compete for acquiring an association with the US and, thus, stimulus competition phenomena such as overshadowing or blocking are viewed as acquisition failures (i.e., the target CS acquires, if anything, a weak association with the US). Thus, in the response rule, negative interaction will occur because the acquisition of  $V_X$  is assumed to take place in a competitive manner.

Second, the response rule allows the target CS,  $X$ , to produce a response based on the associative strength of other CSs,  $i$ , that are activated by the presentation of  $X$ . This assumption of the response rule is not original, but borrowed from the widespread associative view on the role of within-compound associations in higher order learning (for reviews, see Rescorla, 1980; Rescorla & Durlach, 1981). The possibility of allowing a CS to produce responding through the indirect activation of the US representation (i.e., mediated by the activation of other CSs) was previously entertained in the framework of Wagner's (1981) SOP model by Mazur and Wagner (1982) and

was openly adopted by Stout and Miller (2007) in their SOCR model. As in the SOCR model, our response rule explains positive interaction phenomena as being due to the expression of the associative strength gained by those CSs that received compound training with the target CS, X—that is, those CS representations that X activates through a functional within-compound association.

Third, the novelty of X,  $N_X$ , mediates the expression of the associative strength of X's companion CSs,  $i$ . Because the value of  $N_X$  will decrease on each presentation of X, the associative strength of CSs  $i$  will progressively lose impact on the response elicited by X. The underlying assumption here is that responding to a novel CS will depend on the associative strength of both the CS and its associates, whereas responding to a familiar CS will only depend on its own associative strength. That is, as training with the target CS becomes more extensive, responding to this CS will more precisely adjust to its own potential to activate the representation of the US. This progressively weaker influence of X's companion CSs on the response elicited by X as this CS becomes more familiar resembles the well-known transition from stimulus generalization to stimulus discrimination (e.g., Lashley & Wade, 1946; Pavlov, 1927; for recent reviews, see Pearce, 1987, 1994). Although the processes involved in each effect are certainly unrelated, in both cases the response elicited by a novel CS, X, is strongly affected by the associative status of other CSs (i.e., companion CSs in our rule, and physically similar CSs in stimulus generalization). Also, as CS X is further presented, in both cases the response produced by X is gradually less influenced by other CSs (i.e., because of a decrease in X's novelty in our rule, and because of the impact of a differential treatment with the US in stimulus discrimination).

Based on the previous description of the response rule, some important differences emerge with Stout and Miller's (2007) SOCR model: (1) In SOCR, both positive and negative interaction are "pure" performance effects, whereas in our rule positive and negative interaction are assumed to occur at different levels; specifically, negative interaction occurs during acquisition and positive interaction occurs during performance. (2) The modulatory variable of SOCR (the switching operator) mediates in the expression of both positive and negative interaction, whereas in the present rule the modulatory variable (CS novelty) exclusively mediates in the expression of positive interaction. (3) In SOCR, a single type of interaction is expressed at a given time, whereas in the rule proposed here both types of interaction operate simultaneously (assuming  $N_X > 0$ ). In our rule, the balance between positive and negative interaction will determine which one is overtly expressed.

### Some Contrary Associative Effects Explained by the Response Rule

The response rule proposed here can account for a variety of contrary associative effects reported in the literature. The present section describes a series of simulations of some of these effects, which were conducted using the simulation program that can be downloaded from the author's personal website at [www.opineno.com](http://www.opineno.com), as well as from the *Psychonomic Society Archive of Norms, Stimuli,*

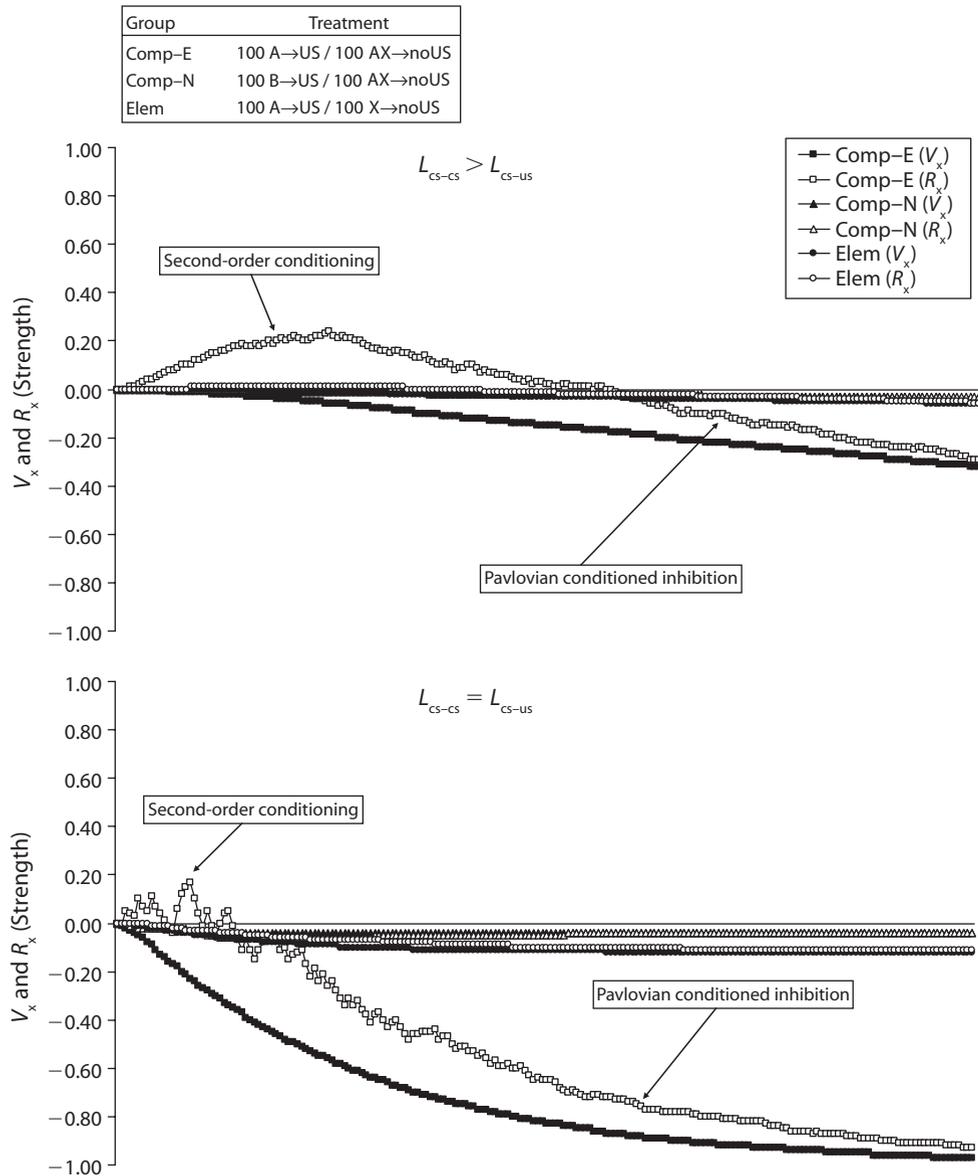
*& Data* at [www.psychonomic.org/archive](http://www.psychonomic.org/archive). Although the present response rule is compatible with any acquisition-focused associative model of learning, for the sake of simplicity and due to its great relevance, the present simulations used the learning rule of the Rescorla and Wagner (1972) model, implemented in the simulation program (see the Appendix for a description of the rules governing the acquisition of CS-US and CS-CS associations in the present simulations). Unless explicitly stated, the simulations employed the default parameters that are set in the simulation program as *predefined values*. Regarding the salience of the stimuli, these parameters are:  $\alpha_A = \alpha_B = 0.50$ ;  $\alpha_X = 0.20$ ,  $\alpha_{\text{contexts}} = 0.10$ ;  $\beta_{\text{US}} = 0.90$ ;  $\beta_{\text{noUS}} = 0.30$ . The value of  $n_{\text{rate}}$  was 0.15. Finally, the default values for the learning-rate parameters were:  $L_{\text{CS-CS}} = 0.80$ ;  $L_{\text{CS-US}} = 0.10$ .

A low salience was selected for the target CS, X, in order to reduce the rate of decrease in its novelty (see Equation 2A), while also reducing X's ability to compete for associative strength. Thus, because of its low salience, CS X could be subject to detectable positive interaction after few training trials (because its novelty decreases slowly) and to detectable negative interaction after extensive training (because other CSs strongly compete with X for the acquisition of associative strength). Also, a low value was chosen for  $L_{\text{CS-US}}$  relative to the value of  $L_{\text{CS-CS}}$  (i.e.,  $L_{\text{CS-CS}} > L_{\text{CS-US}}$ ) in order to slow down the acquisition of interstimulus (CS-US and context-US) associations, therefore allowing within-compound (CS-CS) associations to be acquired relatively faster. This assumption was made based on CS-CS pairings usually being more contiguous and, hence, more effective than CS-US pairings: CS-CS pairings usually involve the simultaneous presentation of the CSs, whereas CS-US pairings usually involve the serial presentation of the CS and the US (i.e., the US is usually presented following the CS presentation). Importantly, this assumption (i.e.,  $L_{\text{CS-CS}} > L_{\text{CS-US}}$ ) will allow the present response rule to simulate positive interaction effects on the early training trials: Because CS-CS associations are acquired more rapidly than CS-US associations, the associative strength of X's companion CSs,  $i$ , will have an impact on responding to X on the initial trials. This point will be illustrated by the simulations in this section, which were conducted in two independent runs employing different sets of values for  $L_{\text{CS-CS}}$  and  $L_{\text{CS-US}}$ :  $L_{\text{CS-CS}} > L_{\text{CS-US}}$  (specifically,  $L_{\text{CS-CS}} = 0.80$  and  $L_{\text{CS-US}} = 0.10$ ) and  $L_{\text{CS-CS}} = L_{\text{CS-US}}$  (specifically,  $L_{\text{CS-CS}} = L_{\text{CS-US}} = 0.80$ ).

The first simulation regards the aforementioned transition from second-order conditioning (e.g., Pavlov, 1927; Rizley & Rescorla, 1972) to Pavlovian conditioned inhibition (e.g., Pavlov, 1927) as a function of the number of trials (Stout et al., 2004; Yin et al., 1994). Figure 1 summarizes the design of the simulation (see table in the inset). In the simulation, condition Comp-E (i.e., Compound-Excitor) received 100 A→US pairings, interspersed with 100 AX→noUS trials. That is, in this condition the target CS, X, was presented in compound with an excitor, A. In control condition Comp-N (i.e., Compound-Neutral), 100 B→US pairings were given, interspersed with

100 AX→noUS trials. Thus, in this condition X was presented in compound with an initially neutral stimulus, A. Finally, condition Elem received 100 A→US pairings, interspersed with 100 X→noUS trials. That is, in this condition X was presented elementally. The panels of Figure 1 depict the results of the simulation, including the associative strength of X ( $V_X$ ) and the strength of the response elicited by X ( $R_X$ ) (the curves reflect simulated test trials

with X alone at every point during training). The top and bottom panels depict the results of the simulation when  $L_{CS-CS} > L_{CS-US}$  and  $L_{CS-CS} = L_{CS-US}$ , respectively. As can be appreciated in both panels, in the control conditions (Comp-N and Elem), X developed a negligible inhibitory association with the US, which also resulted in very weak inhibitory responding. However, in condition Comp-E, X developed an appreciable inhibitory (negative) association



**Figure 1. Second-order conditioning versus Pavlovian conditioned inhibition (the curves reflect simulated test trials with X alone at every point during training). The table (see inset in the figure) summarizes the design of the treatments of this simulation (the letters represent the CSs, “→” means “followed by,” “/” means that the trial types were interspersed, and the numbers represent the numbers of trials with each trial type). The conditions received presentations of X in compound with an excitor (Comp-E), in compound with a neutral stimulus (Comp-N), or elementally (Elem). The simulations in the top and bottom panels only differed in the values adopted by the learning-rate parameters for the CS-CS associations ( $L_{CS-CS}$ ) and the CS-US associations ( $L_{CS-US}$ ). Top panel ( $L_{CS-CS} > L_{CS-US}$ ):  $L_{CS-CS} = 0.80$ ,  $L_{CS-US} = 0.10$ . Bottom panel ( $L_{CS-CS} = L_{CS-US}$ ):  $L_{CS-CS} = L_{CS-US} = 0.80$ . The response produced by X ( $R_X$ ) and the associative strength of X ( $V_X$ ) are depicted by the empty and solid lines, respectively.**

with the US. In the top panel, the X→US inhibitory association developed slowly due to the low value assigned to  $L_{CS-US}$ , and was still incomplete at the end of the simulated treatment (given enough training, this association would fully develop). By contrast, because of the higher value adopted by  $L_{CS-US}$  in the simulation of the bottom panel, this association was acquired relatively faster and reached an asymptotic value by the end of the simulated treatment. Of most importance was that responding elicited by X was excitatory (i.e., indicative of second-order conditioning) after a few training trials and progressively decreased with extended training until becoming inhibitory (by the end of the treatment  $R_X \approx V_X$ ). Although a transitory excitatory response to X can be appreciated in both simulations, this response was stronger and more durable when  $L_{CS-CS} > L_{CS-US}$  (top panel) than when  $L_{CS-CS} = L_{CS-US}$  (bottom panel). That is, in the simulations of Figure 1, second-order conditioning became most evident when the learning-rate parameters were adjusted in order to facilitate the acquisition of CS→CS associations over the acquisition of CS→US associations. However, although weakly, second-order conditioning was also detected in the simulation using identical learning-rate parameters for both CS→CS and CS→US associations. As the following simulation will show, the present response rule more critically depends on the relative values adopted by  $L_{CS-CS}$  and  $L_{CS-US}$  in order to explain other positive interaction phenomena.

The second simulation regards two sets of contrary phenomena that have also been extensively reported in the literature: overshadowing versus potentiation and blocking versus augmentation. *Overshadowing* (e.g., Mackintosh, 1976; Pavlov, 1927) consists of the weak response elicited by a CS, X, that is paired with the US in compound with another CS, A, relative to a condition in which CS X is presented elementally during its pairings with the US. The overshadowing effect is known to be especially strong when the salience of CS X is weak relative to the salience of CS A (i.e., CS A is said to overshadow X based on its higher salience). However, some studies have also reported that, under certain conditions, the presence of A can facilitate, rather than undermine, the acquisition of conditioned responding by X, an effect known as *potentiation* (e.g., Clarke, Westbrook, & Irwin, 1979; Durlach & Rescorla, 1980; Rusiniak, Hankins, Garcia, & Brett, 1979). Analogously, *blocking* (Kamin, 1968) consists of the weak response that is produced by a CS, X, when it is paired with the US in compound with a second CS, A, which was previously paired with the US on its own. In this case, based on its prior training with the US, CS A is said to block the acquisition or expression of conditioned responding by CS X. Interestingly, some studies have found a contrary effect, that is, a facilitative effect of presenting CS A on responding elicited by CS X, a result known as *augmentation* (e.g., Batson & Batsell, 2000).

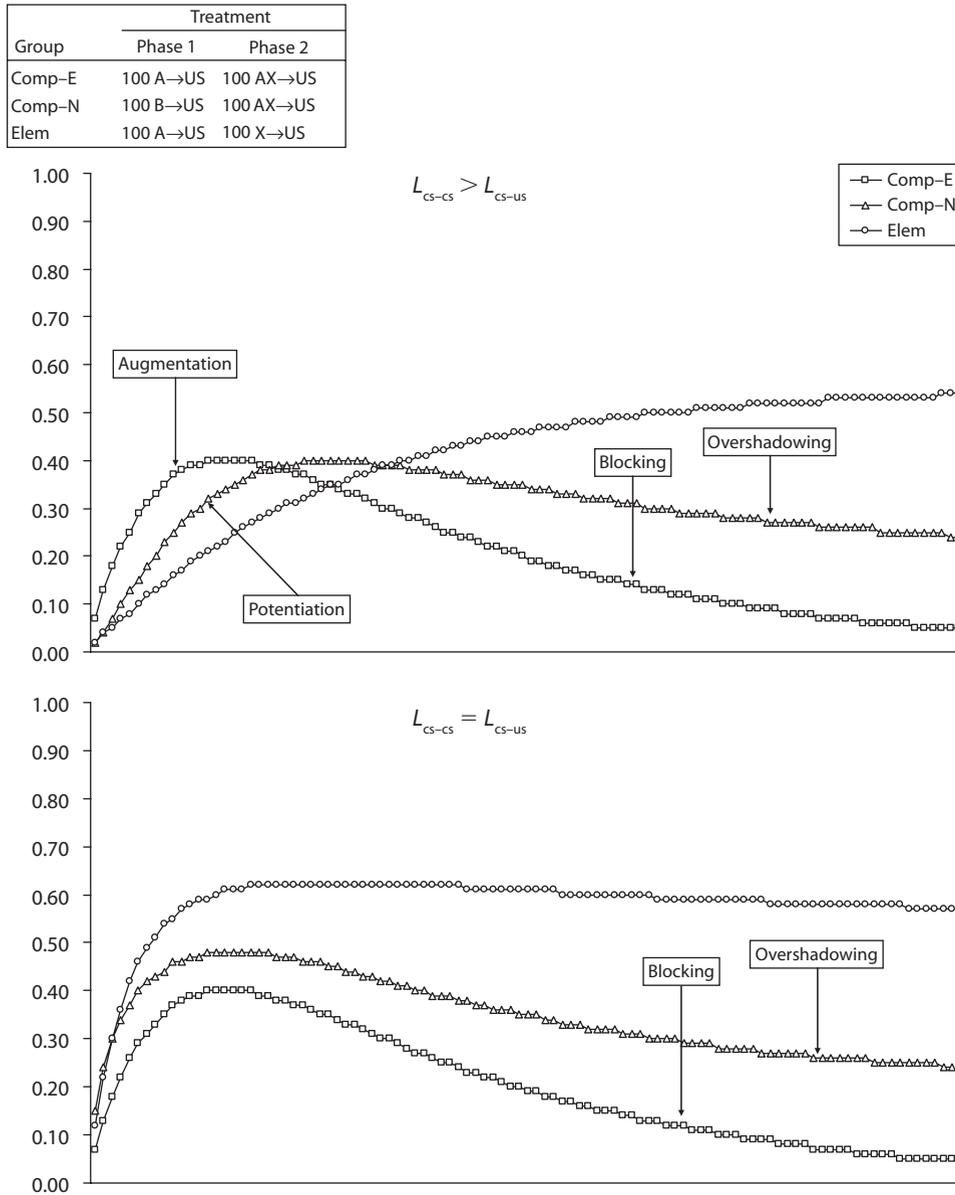
These two sets of contrary effects (i.e., overshadowing vs. potentiation and blocking vs. augmentation) were simulated according to the design of Figure 2 (see table in the inset). The design consisted of three conditions: Condition Comp→E first received A→US pairings in Phase 1, followed by AX→US trials in Phase 2; condition Comp→N

received B→US and AX→US trials in Phases 1 and 2, respectively; and condition Elem received A→US and X→US trials in Phases 1 and 2, respectively. In the present simulation (1) overshadowing would be indicated by weaker responding to X in group Comp→N than in group Elem, whereas the opposite result would be indicative of potentiation, and (2) blocking would be indicated by weaker responding to X in group Comp→E than in groups Comp→N and Elem, whereas the opposite result would be indicative of augmentation. The top and bottom panels of Figure 2 depict the results of the simulation when  $L_{CS-CS} > L_{CS-US}$  and  $L_{CS-CS} = L_{CS-US}$ , respectively. As in Figure 1, the curves reflect simulated test trials with X alone at every point during training (for simplicity, Figure 2 depicts only responding to X during Phase 2). Blocking and overshadowing were observed regardless of the specific learning rate parameters (see top and bottom panels). However, when  $L_{CS-CS} > L_{CS-US}$  (top panel), these effects became evident after a relatively large number of training trials, whereas when  $L_{CS-CS} = L_{CS-US}$  (bottom panel), overshadowing and, specially, blocking were detected from the earliest training trials. In contrast, potentiation and augmentation effects were more elusive: These effects transitorily appeared on the early training trials, and only when  $L_{CS-CS} > L_{CS-US}$  (top panel). That is, as anticipated earlier in this section, potentiation and augmentation were only obtained when the impact of CS A on responding elicited by X on the initial trials was enhanced by assuming that  $L_{CS-CS} > L_{CS-US}$ .

### Some Phenomena the Response Rule Fails to Account For

The previous section showed how the explanatory potential of associative models can be extended to account for the transition from positive to negative interaction as a function of the number of training trials with the target CS. Needless to say, there are many phenomena that this response rule fails to explain. Among these phenomena, perhaps the one posing the most serious challenge to the present response rule regards the fact that negative interaction can be sometimes observed after a single training trial. Evidence of one-trial overshadowing (e.g., Mackintosh, 1971; Revusky, 1971) and one-trial blocking (e.g., Dickinson, Nicholas, & Mackintosh, 1983; Revusky, 1971) is against the predictions of the present response rule since, as shown in the previous section (see simulation in the top panel of Figure 2), this rule predicts that positive interaction phenomena (i.e., potentiation and augmentation) should be observed in this scenario. It is important to mention, however, that this evidence is also problematic for Stout and Miller's (2007) SOCR model. As our response rule, the response rule of the SOCR model necessarily predicts that negative interaction phenomena emerge after an extensive enough treatment with the CSs and, certainly, after more than just one trial.

There are findings that our response rule cannot explain and that are readily accounted for by the SOCR model. For example, evidence showing that overshadowing wanes as the number of AX→US trials is increased (e.g., Stout, Arcediano, Escobar, & Miller, 2003) can be readily ex-



**Figure 2. Overshadowing versus potentiation, and blocking versus augmentation (the curves reflect simulated test trials with X alone at every point during training). The table (see inset) summarizes the design of the treatments of this simulation (the letters represent the CSs, “→” means “followed by,” and the numbers represent numbers of trials with each trial type). The conditions received presentations of X in compound with an excitor (Comp-E), in compound with a neutral cue (Comp-N), or elementally (Elem). The simulations in the top and bottom panels only differed in the values adopted by the learning-rate parameters for the CS–CS associations ( $L_{CS-CS}$ ) and the CS–US associations ( $L_{CS-US}$ ). Top panel ( $L_{CS-CS} > L_{CS-US}$ ):  $L_{CS-CS} = 0.80$ ,  $L_{CS-US} = 0.10$ . Bottom panel ( $L_{CS-CS} = L_{CS-US}$ ):  $L_{CS-CS} = L_{CS-US} = 0.80$ . For simplicity, the figure only depicts responding to X ( $R_X$ ) during Phase 2.**

plained by the SOCR model, while posing a challenge to our response rule. But it is necessary to point out that the SOCR model does not explain such counterintuitive effects based on the variable modulating the occurrence of positive and negative interaction (i.e., the switching operator), but due to its consisting of a mathematical formulation of the extended comparator hypothesis (Denniston

et al., 2001), which already accounted for these effects. As with the waning of overshadowing as a function of the number of AX→US trials, the extended comparator hypothesis is behind the success of the SOCR model in accounting for many phenomena in the literature. Our response rule cannot compete with the SOCR model at this level, but exclusively regarding the transition from posi-

tive to negative interaction as a function of the number of training trials.

Finally, it is also important to point out that the failure of the present response rule in accounting for certain effects is due to our having chosen Rescorla and Wagner's (1972) model as the learning rule in our simulations. As previously mentioned, this response rule is compatible, not only with the Rescorla–Wagner model, but with many other associative models of learning (e.g., Dickinson & Burke, 1996; Mackintosh, 1975; Pearce & Hall, 1980; Van Hamme & Wasserman, 1994; Wagner, 1981). Therefore, some of the boundaries of our response rule in its current version could be easily overcome by merely adopting a more sophisticated learning rule. For example, a relevant effect not explained by the Rescorla–Wagner model is latent inhibition (also known as CS-preexposure effect, Lubow & Moore, 1959), which can be accounted for by a variety of associative models (e.g., Mackintosh, 1975; Pearce & Hall, 1980; Wagner, 1981). Even some effects that, at first sight, could be viewed as belonging to the exclusive domain of the SOCR model, could also be explained if the present response rule employed another learning model. An example is the case of evidence showing that Pavlovian conditioned inhibition (i.e., A→US trials interspersed with AX→noUS trials, see simulations in Figure 1) wanes following extinction treatment with the excitator (i.e., A→noUS trials; e.g., Lysle & Fowler, 1985). This effect can be explained by the SOCR model (again, due to its relying on the extended comparator hypothesis), and could also be explained by our response rule by employing a model that allows learning about absent CSs (e.g., Dickinson & Burke, 1996; Van Hamme & Wasserman, 1994).<sup>2</sup>

In sum, there are a number of phenomena that the present response rule cannot explain. But there are also many phenomena that could be accounted if the learning rule of an associative model different from that of Rescorla and Wagner (1972) was employed. In any case, a fair comparison of the explanatory potentials of our response rule and of Stout and Miller's (2007) SOCR model should exclusively regard their ability to account for the transition from positive to negative interaction. After all, it is only this shift from positive to negative interaction what the present rule attempts to explain.

### Some Effects the Response Rule Can Explain (but SOCR Cannot)

Based on the previous discussion, it is obvious that the present response rule cannot explain all the effects that are readily accounted for by Stout and Miller's (2007) SOCR model. In this section we will discuss some effects that are problematic for the SOCR model, and that our response rule could explain. The SOCR model (as the extended comparator hypothesis, which serves as its basis) has trouble to explain changes in responding to a CS as function of passage of time. Specifically, the SOCR model fails to explain relevant findings in the stimulus competition literature, such as spontaneous recovery of responding after overshadowing (Kraemer, Lariviere, & Spear, 1988) or blocking (Pineño, Urushihara, & Miller,

2005). The problem of the SOCR model is, not only that there is no mechanism by which passage of time could affect responding to a CS, but also that there is no obvious reason why the mechanisms already existing in the model should be affected by passage of time. Although our response rule, as previously described, does not incorporate any mechanism allowing passage of time to have an influence on responding to the target CS, such a mechanism could be easily incorporated in the rule. In order to do so, we could merely assume that the novelty of the target CS, that is, the value of  $N_X$ , increases over time (i.e., provided that the CS is not presented during the retention interval). Such increase in the value of  $N_X$  after a lapse of time could be due to *spontaneous forgetting*.<sup>3</sup> In our response rule, if we assumed that certain features of CS X are forgotten over time, CS X could be processed as a relatively novel stimulus after a relatively long retention interval and, thus, spontaneous recovery of responding after overshadowing and blocking could be explained. This notion of spontaneous forgetting could be introduced in our learning rule as shown by Equation 2C:

$$N_X^t = (N_X^{t-1} + \Delta N_X^t)^{1-SF}. \quad (2C)$$

Equation 2C is identical to Equation 2B, with the only exception that, in Equation 2C, the value of  $N_X^{t-1} + \Delta N_X^t$  maintains an exponential relation with  $1-SF$ . In this equation,  $SF$  denotes the amount of spontaneous forgetting that is caused by the retention interval. The value of  $SF$  would range between 0 and 1, being 0 when no spontaneous forgetting occurs, and 1 when spontaneous forgetting is complete. Therefore, when  $SF = 0$ ,  $N_X^t = N_X^{t-1} + \Delta N_X^t$ , whereas when  $SF = 1$ ,  $N_X^t = 1$ . Thus, as the value of  $SF$  increases toward 1, the value of  $N_X^t$  will accordingly increase toward 1. It must be noted that, in order for response recovery from effects like overshadowing and blocking to be explained, the increase in the value of  $N_X$  should be incomplete: If  $N_X$  adopted a value near 1 (i.e., if the value of  $SF$  was close to 1) following a retention interval, a shift from negative interaction (e.g., overshadowing or blocking) to positive interaction (e.g., potentiation or augmentation) would be predicted. To our knowledge, such a shift has not yet been reported in the literature.

This account of spontaneous recovery from stimulus competition by the present response rule must be considered with caution because an experiment by Pineño et al. (2005, Experiment 3) showed that sensory preconditioning (i.e., a positive interaction effect) waned over a retention interval. This result is problematic for the above suggested possibility that the value of  $N_X$  could increase as a function of passage of time (i.e., because our response rule would expect sensory preconditioning to recover after a retention interval). Pineño et al. suggested that their finding of *spontaneous attenuation* of sensory preconditioning was indicative, at least in their study, of the weakening/forgetting of the within-compound association over time. Thus, it is possible that passage of time causes forgetting of some features of the stimuli previously presented, as well as of the relations previously learned among stimuli. Although speculative, our response rule could reconcile these contrary effects by allowing passage of time

to increase the value of  $N_X$ , while reducing the value of  $V_{X-i}$ . Notably, these contrary effects would arise from a common process, namely, spontaneous forgetting of the CS-features and the within-compound association. If this rather complex approach was followed, a lapse of time could either enhance or impair the expression of positive interaction phenomena depending on the relative impact of spontaneous forgetting on  $N_X$  and  $V_{X-i}$ .

## Conclusion

In this article, a response rule has been proposed to account for positive and negative interaction in stimulus interaction. It has been suggested that this rule can be adopted by associative models (e.g., Rescorla & Wagner, 1972) to explain contrary associative phenomena in the literature. As in Stout and Miller's (2007) SOCR model, the response rule incorporates a variable that modulates the transition from positive to negative interaction as a function of the number of trials with the target CS (e.g., Stout et al., 2004; Yin et al., 1994). In our response rule, this modulatory variable has a well-defined psychological meaning: CS novelty. The present response rule, however, does not necessarily require novelty to be the modulatory variable. Rather, any psychological process that is presumed to change according to experience with the CS could be a potential candidate. For example, this modulatory variable could be *associative priming*. Although somewhat related, priming and novelty greatly differ in the mechanisms producing changes in their values. Novelty is a nonassociative process, and its value decreases merely based on the presentation of the CS. Priming, by contrast, is an associative process, in which the activation of the CS-representation decreases as it is more strongly retrieved or rehearsed (primed) by other CSs or contextual cues. Although both novelty and priming could serve as the modulatory variable in the present response rule, this rule would make different predictions depending on which of these processes was adopted. For example, if the rule included a modulatory process based on priming, a change of context could restore, at least partially, positive interaction. In this case, a CS that is repeatedly presented in the same context would become more strongly primed by the context, thereby reducing the expression of positive interaction. Hence, a change of context should reduce priming and, as a consequence, positive interaction would be restored. A similar effect would be expected to occur if, following repeated presentations of a CS in the same context, the context was extensively presented without the CS. Here, because context-alone presentations would extinguish the previously acquired context-X association, priming of CS X's representation by the context should also decrease, resulting in a shift from negative to positive interaction. Although it can be interesting to entertain the predictions arising from the adoption of one process or another in the modulation of the expression of positive interaction, it is necessary to be cautious until these predictions are exhaustively tested. The use of novelty as the process in the present response rule should be regarded as provisional.

## AUTHOR NOTE

This article was possible thanks to economical support from the Department of Universities, Research, and Technology of the Andalusian Government (Junta de Andalucía). I thank Michael Brown, Ralph R. Miller, Gonzalo P. Urcelay, Miguel A. Vadillo, Jessica M. Zilski-Pineno, and the two anonymous reviewers of a draft of this article for their insightful comments and suggestions. Thanks are also due the Wolk family for their patience and understanding each time I "disappeared" while living at their home in Castleton, VT, in order to work on this article. Correspondence should be addressed to O. Pineño, Department of Psychology, 135 Hofstra University, Hempstead, NY 11550 (e-mail: oskar.pineno@hofstra.edu).

## REFERENCES

- ALLAN, L. G. (1980). A note on measurement of contingency between two binary variables in judgment tasks. *Bulletin of the Psychonomic Society*, **15**, 147-149.
- BATSON, J. D., & BATSELL, W. R. (2000). Augmentation, not blocking, in an A+/AX+ flavor-conditioning procedure. *Psychonomic Bulletin & Review*, **7**, 466-471.
- BRODGEN, W. J. (1939). Sensory preconditioning. *Journal of Experimental Psychology*, **25**, 323-332.
- CHENG, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review*, **104**, 367-405.
- CLARKE, J., WESTBROOK, R. F., & IRWIN, J. (1979). Potentiation instead of overshadowing in the pigeon. *Behavioral & Neural Biology*, **25**, 18-29.
- DENNISTON, J. C., SAVASTANO, H. I., & MILLER, R. R. (2001). The extended comparator hypothesis: Learning by contiguity, responding by relative strength. In R. R. Mowrer & S. B. Klein (Eds.), *Handbook of contemporary learning theories*. Hillsdale, NJ: Erlbaum.
- DICKINSON, A., & BURKE, J. (1996). Within-compound associations mediate the retrospective reevaluation of causality judgements. *Quarterly Journal of Experimental Psychology*, **49B**, 60-80.
- DICKINSON, A., NICHOLAS, D. J., & MACKINTOSH, N. J. (1983). A re-examination of one-trial blocking in conditioned suppression. *Quarterly Journal of Experimental Psychology*, **35B**, 67-79.
- DURLACH, P. J., & RESCORLA, R. A. (1980). Potentiation rather than overshadowing in flavor-aversion learning: An analysis in terms of within-compound associations. *Journal of Experimental Psychology: Animal Behavior Processes*, **6**, 175-187.
- DWYER, D. M. (1999). Retrospective reevaluation or mediated conditioning? The effect of different reinforcers. *Quarterly Journal of Experimental Psychology*, **52B**, 289-306.
- DWYER, D. M. (2001). Mediated conditioning and retrospective reevaluation with LiCl then flavour pairings. *Quarterly Journal of Experimental Psychology*, **54B**, 145-165.
- GEWIRTZ, J. C., & DAVIS, M. (2000). Using Pavlovian higher-order conditioning paradigms to investigate the neural substrates of emotional learning and memory. *Learning & Memory*, **7**, 257-266.
- HOLLAND, P. C. (1981). Acquisition of representation-mediated conditioned food aversions. *Learning & Motivation*, **12**, 1-18.
- HOLLAND, P. C., & FORBES, D. T. (1982). Representation-mediated extinction of conditioned flavor aversions. *Learning & Motivation*, **13**, 454-471.
- KAMIN, L. J. (1968). "Attention-like" processes in classical conditioning. In M. R. Jones (Ed.), *Miami symposium on the prediction of behavior: Aversive stimulation* (pp. 9-31). Miami: University of Miami Press.
- KRAEMER, P. J., LARIVIERE, N. A., & SPEAR, N. E. (1988). Expression of a taste aversion conditioned with an odor-taste compound: Overshadowing is relatively weak in weanlings and decreases over a retention interval in adults. *Animal Learning & Behavior*, **16**, 164-168.
- LASHLEY, K. S., & WADE, M. (1946). The Pavlovian theory of generalization. *Psychological Review*, **53**, 72-87.
- LUBOW, R. E., & MOORE, A. U. (1959). Latent inhibition: The effect of nonreinforced preexposure to the conditioned stimulus. *Journal of Comparative & Physiological Psychology*, **52**, 415-419.
- LYSLE, D. T., & FOWLER, H. (1985). Inhibition as a "slave" process: Deactivation of conditioned inhibition through extinction of conditioned

- excitation. *Journal of Experimental Psychology: Animal Behavior Processes*, **11**, 71-94.
- MACKINTOSH, N. J. (1971). An analysis of overshadowing and blocking. *Quarterly Journal of Experimental Psychology*, **23**, 118-125.
- MACKINTOSH, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review*, **82**, 276-298.
- MACKINTOSH, N. J. (1976). Overshadowing and stimulus intensity. *Animal Learning & Behavior*, **4**, 186-192.
- MAZUR, J. E., & WAGNER, A. R. (1982). An episodic model of associative learning. In M. L. Commons, R. J. Herrnstein, & A. R. Wagner (Eds.), *Quantitative analyses of behavior: Acquisition* (Vol. 3, pp. 3-39). Cambridge, MA: Ballinger.
- MILLER, R. R., & MATZEL, L. D. (1988). The comparator hypothesis: A response rule for the expression of associations. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation* (Vol. 22, pp. 51-92). San Diego: Academic Press.
- PAVLOV, I. P. (1927). *Conditioned reflexes*. London: Clarendon Press.
- PEARCE, J. M. (1987). A model for stimulus generalization in Pavlovian conditioning. *Psychological Review*, **94**, 61-73.
- PEARCE, J. M. (1994). Similarity and discrimination: A selective review and a connectionist model. *Psychological Review*, **101**, 587-607.
- PEARCE, J. M., & HALL, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*, **87**, 532-552.
- PINEÑO, O., URUSHIHARA, K., & MILLER, R. R. (2005). Spontaneous recovery from forward and backward blocking. *Journal of Experimental Psychology: Animal Behavior Processes*, **31**, 172-183.
- RESCORLA, R. A. (1980). *Pavlovian second-order conditioning: Studies in associative learning*. Hillsdale, NJ: Erlbaum.
- RESCORLA, R. A., & DURLACH, P. J. (1981). Within-event learning in Pavlovian conditioning. In N. E. Spear & R. R. Miller (Eds.), *Information processing in animals: Memory mechanisms* (pp. 81-111). Hillsdale, NJ: Erlbaum.
- RESCORLA, R. A., & WAGNER, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64-99). New York: Appleton-Century-Crofts.
- REVUSKY, S. (1971). The role of interference in association over a delay. In W. K. Honig & P. H. R. James (Eds.), *Animal memory*. New York: Academic Press.
- RICCIO, D. C., ACKIL, J., & BURCH-VERNON, A. (1992). Forgetting of stimulus attributes: Methodological implications for assessing associative phenomena. *Psychological Bulletin*, **112**, 433-445.
- RICCIO, D. C., RABINOWITZ, V. C., & AXELROD, S. (1994). Memory: When less is more. *American Psychologist*, **49**, 917-926.
- RICCIO, D. C., RICHARDSON, R., & EBNER, D. L. (1984). Memory retrieval deficits based upon altered contextual cues: A paradox. *Psychological Bulletin*, **96**, 152-165.
- RIZLEY, R. C., & RESCORLA, R. A. (1972). Associations in second-order conditioning and sensory preconditioning. *Journal of Comparative & Physiological Psychology*, **81**, 1-11.
- RUSINIAK, K., HANKINS, W., GARCIA, J., & BRETT, L. (1979). Flavor-illness aversions: Potentiation of odor by taste in rats. *Behavioral & Neural Biology*, **25**, 1-17.
- STOUT, S. [C.], ARCEDIANO, F., ESCOBAR, M., & MILLER, R. R. (2003). Overshadowing as a function of trial number: Dynamics of first- and second-order comparator effects. *Learning & Behavior*, **31**, 85-97.
- STOUT, S. [C.], ESCOBAR, M., & MILLER, R. R. (2004). Trial number and compound stimuli temporal relationship as joint determinants of second-order conditioning and conditioned inhibition. *Learning & Behavior*, **32**, 230-239.
- STOUT, S. C., & MILLER, R. R. (2007). Sometimes competing retrieval (SOCR): A formalization of the comparator hypothesis. *Psychological Review*, **114**, 759-783.
- VAN HAMME, L. J., & WASSERMAN, E. A. (1994). Cue competition in causality judgments: The role of nonpresentation of compound stimulus elements. *Learning & Motivation*, **25**, 127-151.
- WAGNER, A. R. (1981). SOP: A model of automatic memory processing in animal behavior. In N. E. Spear & R. R. Miller (Eds.), *Information processing in animals: Memory mechanisms* (pp. 5-47). Hillsdale, NJ: Erlbaum.
- YIN, H., BARNET, R. C., & MILLER, R. R. (1994). Second-order conditioning and Pavlovian conditioned inhibition: Operational similarities and differences. *Journal of Experimental Psychology: Animal Behavior Processes*, **20**, 419-428.

#### NOTES

1. Another less obvious factor that could be responsible for the observation of either positive or negative interaction is the motivational nature of the US. Dwyer (1999) first considered this possibility based on his finding positive interaction in a conditioned taste preference preparation (i.e., with a high-concentration sucrose solution as the US) and negative interaction in a conditioned taste aversion preparation (i.e., with an emetic, lithium chloride, as the US). However, this possibility was later discarded by Dwyer himself (Dwyer, 2001), who found that the occurrence of each type of interaction depended on a procedural variable, namely the contiguity between the CS and the US.

2. Specifically, in the framework of these models, during the extinction treatment with A, this CS would associatively activate the representations of both CS X and the US and, consequently, an excitatory X-US association would be acquired.

3. Spontaneous forgetting has been observed in the animal conditioning literature as a flattening of the generalization gradients over a period time. That is, stimulus generalization increases after a retention interval, an effect that indicates forgetting of certain features of the CSs as a function of passage of time (see, e.g., Riccio, Ackil, & Burch-Vernon, 1992; Riccio, Rabinowitz, & Axelrod, 1994; Riccio, Richardson, & Ebner, 1984).

**APPENDIX**  
**Learning Rules Employed in the Simulations**

The learning rule for the CS–US associations employed in the present simulations was practically identical to that of the model of Rescorla and Wagner (1972). This rule is represented by Equation A1:

$$\Delta V_{\text{CS-US}}^t = L_{\text{CS-US}} \cdot \alpha \cdot \beta \cdot (\lambda - V_{\text{T-US}}^{t-1}). \quad (\text{A1})$$

In this equation,  $\Delta V_{\text{CS-US}}^t$  represents the change in the strength of a given CS–US association on trial  $t$ .  $\alpha$  and  $\beta$  denote the salience of the CS and the US, respectively ( $0 \leq \alpha \leq 1$ ;  $0 \leq \beta \leq 1$ ). The parenthetical term (i.e.,  $\lambda - V_{\text{T-US}}^{t-1}$ ) represents the discrepancy between the amount of associative strength that can be supported by the US ( $\lambda$ ) and the current total associative strength acquired, until trial  $t-1$ , by all the CSs present on trial  $t$  ( $V_{\text{T-US}}^{t-1}$ ). The value of  $\lambda$  will depend on the presence or absence of the US on trial  $t$  (i.e., US present,  $\lambda = 1$ ; US absent,  $\lambda = 0$ ). Finally,  $L_{\text{CS-US}}$  is a learning-rate parameter for CS–US associations ( $0 \leq L_{\text{CS-US}} \leq 1$ ), which we included in the learning rule of the Rescorla–Wagner model in order to allow CS–US and CS–CS associations to be acquired at different rates.

Equation A2 represents the learning rule for CS–CS associations:

$$\Delta V_{\text{CS}_1\text{-CS}_2}^t = L_{\text{CS-CS}} \cdot \alpha_{\text{CS}_1} \cdot \alpha_{\text{CS}_2} \cdot (1 - V_{\text{CS}_1\text{-CS}_2}^{t-1}). \quad (\text{A2})$$

As can be seen, this rule is very similar to the one employed for learning of CS–US associations. In the present rule,  $\Delta V_{\text{CS}_1\text{-CS}_2}^t$  represents the change in the strength of the association between CS<sub>1</sub> and CS<sub>2</sub>.  $L_{\text{CS-CS}}$  is a learning-rate parameter for CS–CS associations ( $0 \leq L_{\text{CS-CS}} \leq 1$ ), and  $\alpha_{\text{CS}_1}$  and  $\alpha_{\text{CS}_2}$  denote the saliences of CS<sub>1</sub> and CS<sub>2</sub>, respectively. The parenthetical term (i.e.,  $1 - V_{\text{CS}_1\text{-CS}_2}^{t-1}$ ) represents the discrepancy between the maximum strength that can be gained by the CS<sub>1</sub>–CS<sub>2</sub> association (i.e., 1) and the strength acquired, until trial  $t-1$ , by this association ( $V_{\text{CS}_1\text{-CS}_2}^{t-1}$ ).

When either CS<sub>1</sub> or CS<sub>2</sub> is absent on a given trial, the CS<sub>1</sub>–CS<sub>2</sub> association undergoes extinction. Such decrease in the strength of the CS<sub>1</sub>–CS<sub>2</sub> association is governed by a slight variation of Equation A2, represented by Equation A3:

$$\Delta V_{\text{CS}_1\text{-CS}_2}^t = L_{\text{CS-CS}} \cdot \alpha_{\text{CS}_1} \cdot \alpha_{\text{CS}_2} \cdot (0 - V_{\text{CS}_1\text{-CS}_2}^{t-1}). \quad (\text{A3})$$

As can be appreciated, the only difference between Equations A2 and A3 regards the numerical value in the parenthetical term. When either CS<sub>1</sub> or CS<sub>2</sub> are absent on a given trial, the discrepancy resulting from parenthetical term (i.e.,  $0 - V_{\text{CS}_1\text{-CS}_2}^{t-1}$ ) will cause a weakening in the strength of the CS<sub>1</sub>–CS<sub>2</sub> association. In the learning rule for the CS–CS associations represented by Equations A2 and A3, the salience ( $\alpha$ ) of an absent CS is assumed to be lower than that of a present CS (specifically, in our simulations,  $\alpha_{\text{absent CS}} = 1/3 \alpha_{\text{present CS}}$ ). That is, we adopted here an assumption akin to that of the Rescorla–Wagner model on the salience ( $\beta$ ) of present and absent USs. Finally, and also for the sake of consistency with the assumptions of the Rescorla–Wagner model, in our simulations no variation in the strength of the CS<sub>1</sub>–CS<sub>2</sub> association was presumed to occur when both CS<sub>1</sub> and CS<sub>2</sub> are absent.