In S. Vosniadou, D. Kayser, & A. Protopapas (Eds.) (2007). Proceedings of the European Cognitive Science Conference, EuroCogSci07 (pp. 341-346). Hove, UK: Erlbaum.

Either greedy or well informed: The reward maximization – unbiased evaluation trade-off

Helena Matute (matute@fice.deusto.es)
Miguel A. Vadillo (mvadillo@fice.deusto.es)
Fernando Blanco (fblanco@fice.deusto.es)
Serban C. Musca (serbancmusca@gmail.com)

Departamento de Psicología, Universidad de Deusto
Apartado 1, 48080 Bilbao, SPAIN

Abstract

People often believe that they exert control on uncontrollable outcomes, a phenomenon that has been called illusion of control. Psychologists tend to attribute this illusion to personality variables. However, we present simulations showing that the illusion of control can be explained at a simpler level of analysis. In brief, if a person desires an outcome and tends to act as often as possible in order to get it, this person will never be able to know that the outcome could have occurred with the same probability if he/she had done nothing. Our simulations show that a very high probability of action is usually the best possible strategy if one wants to maximize the likelihood of occurrence of a desired event, but the choice of this strategy gives rise to illusion of control.

Introduction

The illusion of control has been observed in many different laboratory experiments since the initial studies by Langer (1975). It consists of people believing that they have control over desired outcomes that are uncontrollable but occur frequently. As a real life example, let us think of the way ancient tribes danced for rain, or the way many people, still today, believe in magical rituals rather than in scientific medicine as the best means to improve their health. These examples should give us an idea of the prevalence and importance of this problem in relation to human welfare.

Most explanations for this effect have been framed in terms of personality and self-esteem protection (e.g., Alloy & Abramson, 1982). However, and without discussing the importance of personality variables, what we would like to argue is that the basic tendency towards an illusion of control is present in all of us, as it is just a consequence of the way we interact with the world when we want to influence the occurrence of events. We will make use of simulations to illustrate our point.

The basic idea is a very simple one. Imagine a person who is trying to obtain an outcome that is of crucial importance for survival. Quite probably, this person will tend to act at every opportunity in order to obtain it. If the outcome is uncontrollable but occurs frequently, if this person is responding as often as possible, the occurrence of the outcome will surely coincide with the person's action most of the time. Thus, it is not strange that under such conditions, this person will develop an illusion of control. In order to be able to realize that the outcome would have occurred with the same probability regardless of responding, this person should adopt a much more scientific strategy: he or she should test not only what happens when a response is performed but also what happens when a response is not performed. That is, they should respond only in 50% of the trials so that they can equally sample both cases. However, are people ready to test what happens in the absence of a magical ritual when they believe that the ritual is responsible for a very important outcome?

The many studies that have been published showing that laboratory participants are indeed able to detect when outcomes are uncontrollable (e.g., Shanks & Dickinson, 1987; Wasserman, 1990) would make us believe that people do naturally behave in the scientific way described above and naturally detect response-outcome contingencies. However, those laboratory studies instruct their subjects very explicitly on how to behave and what to look for. If we manipulate the instructions that participants receive in an uncontrollable situation, participants who are simply instructed to obtain the outcomes tend to respond at every opportunity (and therefore, to develop an illusion of control as well); on the other hand, those participants who are instructed to adopt the scientific strategy, are the ones who are able to realize that the task is uncontrollable (Matute, 1996). In other words, people do have the cognitive capacity to detect the absence of control, but this does not necessarily mean that they will use it by default, in naturalistic settings. Indeed, Matute's (1996) studies suggested that, unless there is a special motivation to detect the degree of control that one has over the outcome, people will tend to respond as much as possible, rather than in 50% of the trials. In the present research we will show that even for an artificial system, responding as much as possible is the best possible strategy when its aim is to obtain an outcome that is controllable; but the counterpart of behaving this way is that the system will be more prone to develop an illusion of control when faced with uncontrollable situations.

Simulations

Procedure

Our simulations are based on the Rescorla-Wagner model (Rescorla & Wagner, 1972) model, which is one of the most widely used in the area of learning research to simulate how people learn to associate potential causes and effects (like, for example, responses and outcomes). This model is formally equivalent to the delta rule (Widrow & Hoff, 1960) used to train two-layer distributed neural networks through a gradient descent learning procedure. In the Rescorla-Wagner model the change (ΔV_R^n) in the strength of the association between a potential cause (in our case, the system's response, R) and a potential effect (a desired outcome) after each learning trial, takes place according to the following equation:

$$\Delta V_R^n = k \cdot (\lambda - V_t^{n-1}) \quad (1)$$

where k is a learning rate parameter that reflect the associability of the cause, α, and that of the effect, β, ($K = \alpha \cdot \beta$ in the original Rescorla & Wagner model); λ reflects the asymptote of the curve (which is assumed to be 1 in trials in which the outcome is present and 0 otherwise), and V_t^{n-1} is the strength with which the effect can be predicted by the sum of the strengths that all the possible causes that are present in the current trial had in trial *n-1*. For example, in a simulation of the illusion of control, there should be at least two possible causes for the occurrence of the outcome: one is the system's response, R, the other one is the context in which the response takes place (see, e.g., Shanks & Dickinson, 1987). Thus, for instance, when the outcome occurs but there is no response, the occurrence of the outcome will be attributed to other, background or contextual, potential causes. By the same reasoning, when the outcome occurs after a response has been given, the outcome will be attributed to both the response and the context, as a function of their respective associability. The task of the learner will be to learn how much is due to his or her own response, how much is due to other, unspecified potential causes. In general, contexts are assumed to be of low associability, thus, in all the simulations that we will report, k will be 0.10 for the context and 0.30 for the response. Also, it is often the case in many published simulations of this model that k takes different values as a function of whether the outcome occurs or as a function of age-related or species-related differences in sensitivity to the outcome. However, for the sake of simplicity we have preferred to ignore these additional parameters in our simulations. Thus, the value of k, for both the context and the response, will be kept constant, regardless of whether the outcome occurs or not. For each simulation, 100 learning trials and 500 iterations will be run.

In all simulations, the probability that the outcome occurs when the system makes a response, p(O|R), will be 0.75. The probability that the outcome occurs when there is no response, p(O|noR), will be 0.75 in some simulations and 0 in others. When those two probabilities are identical (e.g., both of them are 0.75), the outcome is said to be noncontingent on the response, or, in other words, *uncontrollable*. In this case, the actual contingency is 0 (i.e., 0.75 – 0.75). When these two probabilities are different (i.e., 0.75 and 0, respectively), then the outcome is controllable

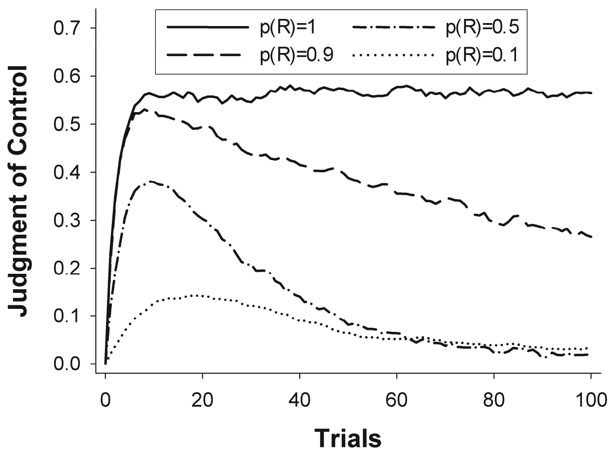

Figure 1: In Simulation 1 outcomes occur with a probability of 0.75 and are uncontrollable (i.e., they occur regardless of whether the system responds or not). The judgment of control is shown to depend on the probability of responding. (See main text for simulation details.)

(i.e., there is a positive contingency of 0.75). Thus, we will test both controllable and uncontrollable conditions. The reason why we are using a high probability of the outcome's occurrence (i.e., 0.75) both in controllable and uncontrollable conditions is that the illusion of control is more readily observed in uncontrollable conditions when the outcome occurs frequently (e.g., Alloy & Abramson, 1979; Matute, 1995).

The strength of the association between the response and the outcome is taken as an index of the strength of the response-outcome causal relation perceived by the system (i.e., the judgment of control). Thus, an illusion of control will be observed anytime when the strength of the association between the response and the outcome becomes higher than zero in a noncontingent situation.

Across simulations we will manipulate the probability that the system responds in each trial, p(R). In the first set of simulations we will compare the effect of different probabilities of responding, ranging from 0.1 to 1.0. In the second set of simulations, probabilities of responding will not be fixed, as they will change with experience.

Results

Simulations using a fixed p(R) Simulation 1 considers a noncontingent situation where the outcome occurs in 75% of the trials, regardless of whether there is a response or not. The results of this simulation, presented in Figure 1, show that the illusion of control is dependent on the probability of responding: As the probability of acting approaches 1, the illusion of control becomes stronger and more persistent over trials.

Now, if responding with a very high probability produces such illusions, why do people tend to respond so much? Wouldn't it make more sense to be less active so that the

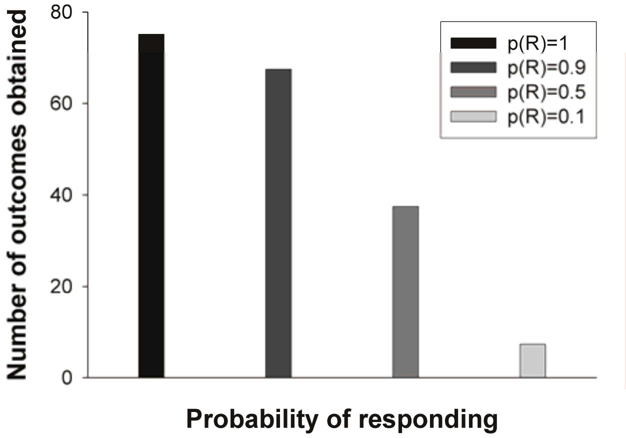

Figure 2. In Simulation 2 the outcome is said to be controllable because it occurs in 75% of the occasions in which the system responds and never in its absence. Simulation 2 shows that the number of outcomes that is obtained after 100 trials is considerably reduced as the probability of responding departs from 1.

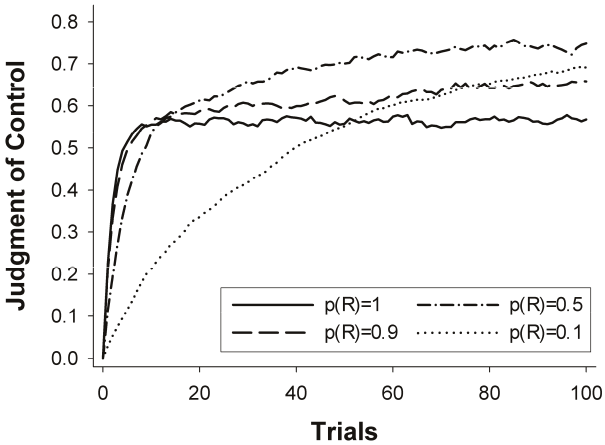

Figure 3. Simulation 3 uses the same controllable condition as Simulation 2 (i.e., the outcome occurs in 75% of the occasions in which the subject responds and never in the absence of responding), but here the dependent variable is the judgment of control (associative strength). Simulation 3 shows that, even in contingent conditions, the high p(R) strategy is not the best one with respect to contingency detection.

actual contingency could be accurately detected? If a system is trying to find out how much control is available over an uncontrollable outcome, this system should, as shown in Simulation 1, be quite passive. A low probability of responding will certainly allow the system to correctly detect the uncontrollability of the outcome and would not affect the amount of the outcomes obtained, since in uncontrollable situations responding with a high or low probability does not affect the amount of outcomes that can be obtained.

However, let us now imagine a situation in which the outcome effectively depends on the subject's behavior. Thus, in Simulation 2, the outcome *is* controllable. Assume, for example, that the outcome occurs in 75% of the occasions in which the system responds, and it never occurs when the system does not respond. This case is shown in Figure 2: A system that acts with a probability of 1 will be able to obtain more outcomes than a system responding with at a lower probability. As the probability of responding drops down from 1, the percentage of desired outcomes obtained is reduced. This, of course, is true for any positive contingency situation (and the opposite is true for negative contingency). Thus, for any condition that depends on our performing a given action, the best thing we can do in order to maximize reward is to perform the action just in all occasions (Simulation 2). The bad news is that this strategy will produce an illusion of control when the outcome is uncontrollable (Simulation 1).

It is clear that the best strategy to maximize the number of outcomes are not optimal when the goal is to know how much control one has over the outcome. If the outcome happens to be uncontrollable, the high p(R) strategy will provide the user with data that is too noisy and incomplete to accurately calculate the actual contingency, thus giving rise to illusion of control. But, is the high p(R) strategy problematic only in noncontingent situations?

Simulation 3 compares the detection of contingency that can take place in a contingent situation when the probability of responding is 1 as compared to when it is reduced (up to 0.1). Simulation 3 was conducted in the same conditions as Simulation 2, but the dependent variable is now the strength of the association (or judgment of control) rather than the number of outcomes obtained. Thus, it considers a contingent relation in which the outcome occurs in 75% of the trials in which the subject responds and never when there is no response. As can be seen in Figure 3, even when the outcome is contingent on responding – and therefore, the best thing one can do to maximize reward is to respond in all occasions (cf. Simulation 2) – the high p(R) strategy prevents the accurate detection of the contingency. In this case, the actual contingency is 0.75. Even a subject responding with a very low probability (0.1) will be able to produce a much better judgment of control than one who responds always. In this later case, there is no illusion in our high p(R) system because the outcome is contingent on responding, but the contingency that this system perceives between the response and the outcome is lower than the one that is actually present.

This may seem surprising at first. However, as was the case in the noncontingent conditions shown in Simulation 1, if the system responds in every single trial, it cannot know what happens when there is no response. In this case, the subject is just exposed to what happens when the response is given in a given context. And, according to Equation 1, the increment in the strength of the association that can be accrued in a given trial depends not only on the strength of the association between the response and the outcome in the

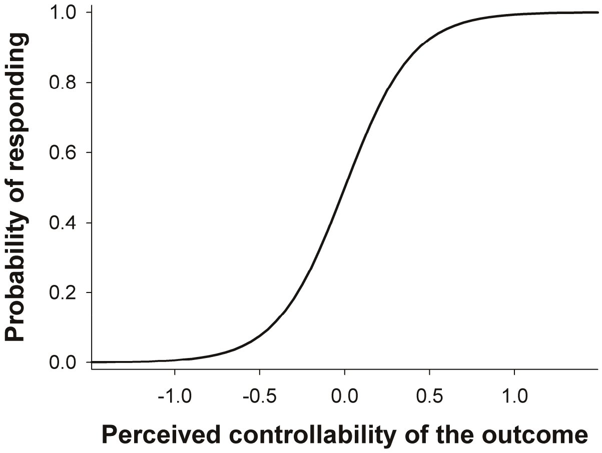

Figure 4. Sigmoid function for the probability of responding based on the perceived controllability of the outcome (i.e., on the strength of the response-outcome association).

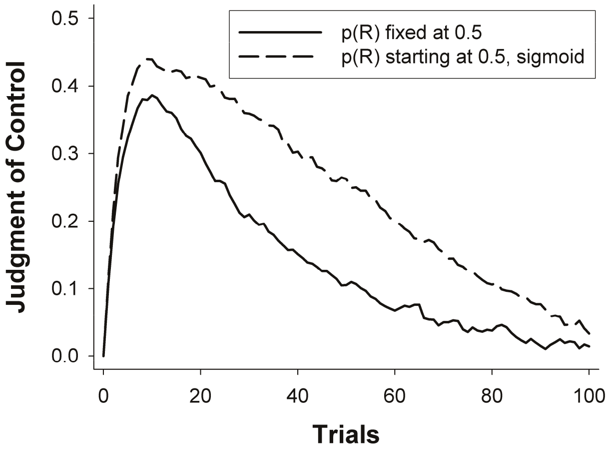

Figure 5. In Simulation 4 uncontrollable outcomes occur with a probability of 0.75. The illusion of control is more intense and persistent when the system's p(R) varies according to the strength of the response than when this probability is fixed.

previous trial, but also on the strength with which the other cues that are present (e.g., the context) are already associated with the outcome. This means that the associative strength that could be accrued by the response in a given trial will be shared by the response and the context (as a function of their relative associability; *k* in the equation, and their associative strength in the previous trial; V_t^{n-1} in the equation). By the same reasoning, the trials in which the response does not occur (in systems in which the p(R) is different from 1), can only affect the strength of the context. And, because in the contingent situation we are testing in Simulation 3, the outcome does not occur when there is no response, these no-response trials will reduce the strength of the context alone. Moreover, the reduction of the strength of the context will in turn have the (indirect) effect of increasing the strength of the response. This is because, after the context strength has been reduced, when a response is given in a subsequent trial, the competition that the context can exert for associative strength will be lower. In this way the response will get a larger proportion of the available strength in all systems responding with a p(R) lower than 1 in Simulation 3 (see Equation 1). However, a system that responds with a p(R) of 1 does not have information on what happens when the response is absent and just the context is present. In other words, there are no context-alone trials that will help the system discard the potential causal role of the context. If this is so, then the associative strength accrued by the response *and* the context in each trial (appreciate that they always occur in compound in this system) are shared between the two of them as a function of their respective *k*s. This is why it is impossible for a subject responding at every opportunity to accurately detect contingencies, not only in uncontrollable situations but also in controllable ones. As shown in Simulation 3, a subject responding with a probability of 0.9, or even 0.1 will be much more accurate in the detection of the actual contingency than a subject responding all the time, even when the outcome is controllable. Still, one has to keep in mind that these would not be good strategies if what we want is to maximize reward.

Simulations using a modifiable p(R) One could argue that our previous simulations use artificial conditions, in that living organisms do not keep a fixed probability of responding regardless of what they learn; by contrast, they vary their probability of responding as a function of how strongly they believe that the response is the cause of the outcome. Thus, let us now suppose that if a response is very strongly associated to the outcome (in other words, the system believes the response is the cause of the outcome), the probability of responding will be stronger.

Simulation 4 (see Figure 5) is similar to the previous ones, but here the probability of responding is increased or reduced as a function of the strength of the association that is being learned. To this end, we use a simple sigmoid function that increases the probability of responding when the association increases and reduces it otherwise:

$$p(R) = 1/(1+e^{-\theta \cdot V_R^{n-1}}) \quad (2)$$

For the present purposes, the parameter describing the slope of the sigmoid function, θ, was set to 5. Figure 4 depicts the different values that p(R) can receive depending on the strength of the response-outcome association. As there can be seen, a system acting according to this equation will simply tend to respond with a very high probability when the response is apparently causing the outcome. If the perceived contingency between the response and the outcome is negative (that is, if the system believes that the response actually prevents the occurrence of the outcome), the probability of responding would be near 0. Finally, when the associative strength is near 0 and, therefore, the system believes that the outcomes are uncontrollable, the probability of response is intermediate.

Note that in Equation 2 the probability of responding is dependent on the strength of the association. This implies that for the first trial the probability of responding is to some extent arbitrary, because for the first trial there is no prior associative strength upon which to compute the probability of responding. In Simulation 4, the probability of responding for the first trial was set to the intermediate value of 0.50.

Thus, Simulation 4 corresponds to a more natural condition than the previous ones, in that cognitive systems generally vary their probability of responding according to the strength of the association that they have formed between the response and the outcome (or, in other words, the strength that they attribute to their own response as a cause of the outcome). As can be seen in Figure 5, the illusion of control that is developed in this way is even more intense and persistent than the one produced by a fixed p(R), as that used in Simulation 1.

But let us now suppose that not only do subjects vary their probability of responding as a function of what they learn, but also that different subjects probably start up from different backgrounds, beliefs, strategies… and personalities. This should at least produce some initial biases. These differences in the initial conditions, even though they are subsequently subject to a common learning function that will tend to make them similar to each other at asymptote, could perhaps produce important differences in the speed and slope of learning.

Simulation 5 tests whether the apparently innocuous little biases that many people may have during the initial stages of a new task (e.g. being more of less active), can have a profound effect on the strength and the durability of the illusion of control. This simulation is very similar to Simulation 4, but here two systems that are sensitive to the strength of the association (i.e., that use a sigmoid function, as in Simulation 4) are compared. The probability of making a response in the very first trial is what we manipulated here. The difference between the two systems is that the probability of responding in the first trial is 0.1 for one of them and 0.9 for the other. In all remaining trials, the probability of responding in both systems is computed according to Equation 2.

The results are presented in Figure 6. The initial bias – that represent the tendency to respond more or less due to previous history, background, beliefs, or personality – though implemented only in the very first trial still has an effect after 100 trials.

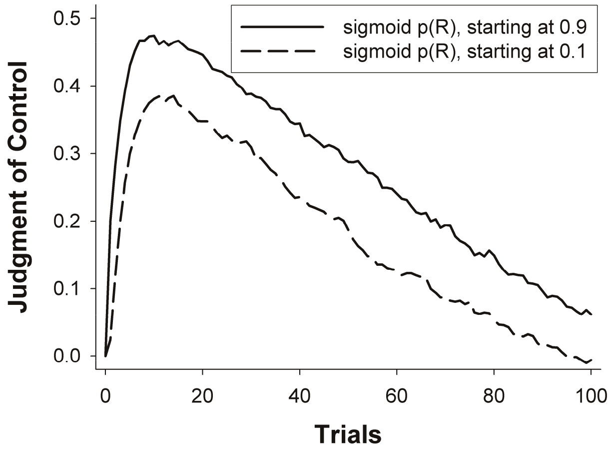

Figure 6. In Simulation 5 uncontrollable outcomes occur with a probability of 0.75 regardless of whether the system responds or not, as in Simulation 4. The two systems here considered do vary their probability of responding according to the strength of the association between the response and the outcome, but one of them starts with a stronger bias to act in the very first trial. This initial, first trial bias still has an effect on performance after 100 trials.

Discussion

The illusion of control is at the roots of many real world problems, like the reluctance of many people to believe in scientific medicine and the proliferation in today's world of so many magical and pseudoscientific remedies for almost everything. It is generally believed to be part of naïve personalities, but we have shown that it is potentially a much more prevalent problem that can occur in all cognitive systems. Indeed, it is a logical consequence of how we interact with the world. Even though personality variables can also have an important influence and can surely be responsible for individual differences among people, they are not the only variables that are responsible, nor the only ones that should be taken into account when trying to set therapies and policies to eradicate this illusion. As shown in our simulations, the main problem has to do with what the goal of the system is. If our goal in the world is to maximize the number of rewards (and this is an important goal for survival that can certainly have been favored by evolution as an adequate strategy for many occasions), then the system will try to respond as much as possible in order to obtain those outcomes. As shown in Simulation 2, only those subjects responding in all possible occasions will get the majority of the available rewards when the situation is controllable (of course, this would be irrelevant if the situation were noncontingent). Thus it would not be strange that a default strategy in many people and even in animals would be to respond as much as possible. What is clear from our simulations is that this strategy, while optimal when one wants to maximize reward, is quite a bad one in the occasions in which the goal of the system is not to obtain the outcome, but to analyze to what degree it is controllable. Therefore, it is to some extent contradictory trying to maximize control over the environment and, at the same time, trying to make accurate inferences about the world. This means that, if a given outcome is important enough for people, the attempts they make to control it will surely interfere with the ability to accurately assess the degree of control they actually have.

In sum, imagine that twenty people were suddenly infected with an unknown mortal disease and that, for some reason, you suspect that medicine X might cure them. Would you be ready to test this medicine just in one half of your patients so as to check that the medicine is actually working? This is actually the difference between scientific reasoning and every day reasoning. As we have shown, none of these strategies can be said to be better than the other one; it is only a matter of choosing the right one at the right time. Thus if we would like people to apply more scientific reasoning to their everyday life, perhaps we should start by trying to convince them to test passive responding in situations in which the outcome is unimportant for them. In this way, they will be able to learn what they need about skepticism so that the next time they face a serious problem they will be able to actively chose the p(R) strategy that best complies with their own goals.

Acknowledgments

Support for this research was provided by Grant SEJ406 from Junta de Andalucía. Fernando Blanco was supported by a F.P.I. fellowship from Gobierno Vasco (Ref.: BFI04.484). We would like to thank Cristina Orgaz for valuable discussions on these points. Correspondence concerning this article should be addressed to Helena Matute, Departamento de Psicología, Universidad de Deusto, Apartado 1, 48080 Bilbao, Spain. E-mail: matute@fice.deusto.es.

References

Alloy, L. B., & Abramson, L. Y. (1979). Judgment of contingency in depressed and nondepressed students: Sadder but wiser? *Journal of Experimental Psychology: General, 108,* 441-485.

Alloy, L. B., & Abramson, L. Y. (1982). Learned helplessness, depression, and the illusion of control. *Journal of Personality and Social Psychology, 42,* 1114-1126.

Langer, E. J. (1975). The illusion of control. *Journal of Personality and Social Psychology, 32*, 311-328.

Matute, H. (1995). Human reactions to uncontrollable outcomes: Further evidence for superstitions rather than helplessness. *Quarterly Journal of Experimental Psychology, 48B,* 142-157

Matute, H. (1996). Illusion of control: Detecting response-outcome independence in analytic but not in naturalistic conditions. *Psychological Science, 7,* 289-293.

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: Appelton-Century-Crofts.

Shanks, D. R., & Dickinson, A. (1987). Associative accounts of causality judgment. In G. H. Bower (Ed.), *The psychology of learning and motivation, Vol. 21* (pp. 229-261). San Diego, CA: Academic Press.

Wasserman, E. A. (1990). Detecting response-outcome relations: Toward an understanding of the causal texture of the environment. In G. H. Bower (Ed.), *The psychology of learning and motivation, Vol. 26* (pp. 27-82). San Diego, CA: Academic Press.

Widrow, B., & Hoff, M. E. (1960). Adaptive switching circuits. *1960 IRE WESCON Convention Record* (pp. 96-104). New York: IRE.