Running Head: DISSOCIATIONS AND COGNITIVE PRIORITY

IN PRESS: ***Canadian Journal of Experimental Psychology***

Available at http://www.apa.org/pubs/journals/cep/index.aspx. This article may not exactly replicate the final version published in the CPA journal. It is not the copy or record.

Dissociations among judgments do not reflect cognitive priority: An associative explanation of memory for frequency information in contingency learning

Miguel A. Vadillo[1] & David Luque[2]

[1]*Universidad de Deusto, Bilbao, Spain*
[2]*Universidad de Málaga, Málaga, Spain*

Mailing address:
Miguel A. Vadillo
Dpto. de Fundamentos y Métodos de la Psicología
Universidad de Deusto
Apartado 1, 48080 Bilbao, Spain

Fax: + 34 – 94 – 4139089

e-mail: mvadillo@deusto.es

Abstract

Previous research on causal learning has usually made strong claims about the relative complexity and temporal priority of some processes over others based on evidence about dissociations between several types of judgments. Specifically, it has been argued that the dissociation between causal judgments and trial-type frequency information is incompatible with the general cognitive architecture proposed by associative models. In contrast with this view, we conduct an associative analysis of this process showing that this need not be the case. We conclude that any attempt to gain a better insight on the cognitive architecture involved in contingency learning cannot rely solely on data about these dissociations.

Keywords: Contingency learning; probability learning; statistical models; associative models.

Inferring the causal structure of the environment is an invaluable skill for survival in an ever changing environment. As first noted by Hume (1739/1964), we are unable to directly perceive the link connecting causes and effects, which means that causal relations can only be inferred on the basis of indirect evidence. Unfortunately, it is not easy to determine when a given event is a real cause of an effect, because some events can co-occur regularly without any causal link between them. Psychological models of causal learning try to explain *when* and *how* humans can learn new causal links.

Following the advice of Marr (1982), some authors have focused on developing computational models of causal learning (e.g., Allan, 1980; Cheng, 1997; Cheng & Novick, 1992; Holyoak & Cheng, 2011; Pearl, 2000). According to the popular levels-of-processing framework, computational models do not aim at specifying every step that the cognitive system has to give in order to solve a problem. Instead, these models should clarify, for a given task, what is the function that maps the input of the cognitive system with its output, while being agnostic about the algorithm involved in that computation. Following this general perspective, several *statistical models* of causal learning have been proposed to describe what input allows us to determine that two events are causally related. These models usually consist of simple mathematical equations that provide a numerical index of the strength of the relationship between a candidate cause and an effect: Values different from zero usually indicate that a causal relation exists. Note that, in principle, computational models are not concerned about *how* humans actually acquire this causal knowledge: Computational models only establish *when* a causal link must be inferred (if the system works well).

Interestingly, some authors have proposed that the statistical calculations proposed by these computational models could also be appropriate theories about *how* humans actually learn new causal associations. According to these authors, the mathematical

formulas of *statistical models* of causal learning can also be considered algorithm-level theories of causal learning that specify the real steps that people give when they solve a causal induction problem. From this algorithm-level viewpoint, people act as intuitive statisticians who first gather information about the joint occurrence of two events and then combine this information following certain rules in order to decide whether or not there is an statistical connection between those two events. For instance, when faced with a sequence of trials in which a cause, C, and an effect, E, appear together or in isolation, people are assumed to first encode this information in a mental representation to some extent isomorphic to the 2 x 2 contingency table depicted in Table 1 (Beyth-Marom, 1982; Busemeyer, 1991; Shaklee & Mims, 1982). This contingency table would summarize the evidence experienced by the participant regarding the joint occurrence or absence of the target cause and effect, including the number of occasions in which both the cause and the effect have appeared together (*a*), the number of occasions in which only the effect or only the cause has appeared (*b* or *c*, respectively), and the number of occasions in which both the cause and the effect have been simultaneously absent (*d*). Once the participant has stored this information, he or she can use these frequency data to compute a contingency index that describes the strength of the causal link between the cause and the effect. For example, it has been argued that one of the contingency indexes that participants might try to compute from frequency information is the so-called Δp index, which is defined as the extent to which the occurrence of the cause increases the probability of the effect (Allan, 1980; Jenkins & Ward, 1965). As can be seen in the following equation, this index can be easily computed from the data contained in the 2 x 2 contingency table described in Table 1:

$$\Delta p = p(\mathrm{E}|\mathrm{C}) - p(\mathrm{E}|\sim\mathrm{C}) = a / (a + b) - c / (c + d) \qquad \text{(Equation 1)}$$

Many experiments have shown that people judge causal strengths in a way consistent with the predictions of the Δ*p* rule (e.g., Shanks, 1987; Shanks & Dickinson, 1987, 1991; Wasserman, 1990; Wasserman, Elek, Chatlosh, & Baker, 1993). Moreover, although it has sometimes been found that peoples' judgments deviate from the Δ*p* rule in certain conditions (e.g., Allan & Jenkins, 1980; Alloy & Abramson, 1979; Jenkins & Ward, 1965; Smedslund, 1963), it is usually easy to propose alternative statistical rules that participants could be applying on frequency data that could potentially account for these deviations (Allan & Jenkins, 1980, 1983; Busemeyer, 1991; White, 2003).

In contrast with this general framework, some researchers have proposed an alternative algorithm-level account of contingency detection. From this alternative approach, contingency detection, and causal learning in general, is just another instance of associative learning, comparable to simpler processes such as classical or instrumental conditioning (Alloy & Abramson, 1979; Dickinson, Shanks, & Evenden, 1984). In other words, people would associate causes and effects just the same way they associate, in general, any type of cues and outcomes that are regularly paired in the environment. The key assumption of these models is that while participants are exposed to a sequence of cause-effect pairings (or cue-outcome pairings, in the more neutral, associative terminology), an association between the mental representations of the cause and the effect is formed. The strength of the cause-effect association is assumed to change on a trial-by-trial basis according to a simple error-correction algorithm similar to the ones used in connectionist simulations of cognitive processes. One of the most widely used algorithms used to model contingency learning is the rule initially proposed by Rescorla and Wagner (1972) in the domain of classical conditioning (RW rule, henceforth). According to this model, in any given trial, the strength of a cause-effect association, $\Delta V_{C\text{-}E}$, changes according to the following equation:

$$\Delta V_{C\text{-}E} = \alpha \cdot \beta \cdot (\lambda - V_{TOTAL}) \qquad \text{(Equation 2)}$$

where λ denotes the presence or absence of the effect in that trial (usually coded as 1 or 0, respectively), V_{TOTAL} is the sum of the associative strengths of all the cues presented on that trial (usually including not only the strength of the target cause, $V_{C\text{-}E}$, but also that of a constant context, $V_{CTX\text{-}E}$) and α and β are learning-rate parameters dependent on the salience of the cause and the effect, respectively, with values ranging from 0 to 1. The associative strength of the context would also change trial-by-trial according to the same equation, although with a different α that would represent the salience of the context. The associative strength of the context would be updated even in trials in which the target cause is nevertheless absent.

Many of the experiments conducted during the last two decades aimed at discriminating between rule-based and associative accounts of causal learning (Allan, 1993, 2003). However, this task is more difficult than it could be expected *prima facie*, as the predictions made by both accounts are virtually indistinguishable under many conditions. For example, it is well known that in situations in which a single cue-outcome association is being learned, the asymptotic strength of a cue-outcome association as computed with the RW rule converges to the value of the contingency as computed with the Δ*p* rule (Chapman & Robbins, 1990; Wasserman, Elek, Chatlosh, & Baker, 1993; see also Danks, 2003).[1] Therefore, any evidence showing that participants' judgments of a causal relation fit with the predictions of the Δ*p* rule is usually also explainable in terms of the RW learning algorithm. More generally, both kinds of theories share important assumptions: They assume a similar representation format (a single scalar parameter describing the strength of the cause-effect relationship) and also a similar normative analysis of what causal induction is. The main difference between these models lays in the algorithmic details specifying how these computations are carried out. In the case of rule-

based models, people are assumed to store the frequency information and use that information to compute an index of contingency when asked to do so, whereas in associative models they are assumed to constantly update the strength of the cause-effect association on a trial-by-trial basis by means of an error-correction mechanism. Although the algorithmic details and the cognitive architecture assumed by both theories are very different, the final output the process can be remarkably similar.

Given the similarities between the predictions made by rule-based and associative accounts of causal learning, some studies have attempted to test them not by contrasting their differential predictions, but by checking the plausibility of the different mechanisms invoked by both types of models. Specifically, researchers have relied on at least three types of evidence to draw conclusions about the relative success of rule-based modes over associative models.

First, in contrast to associative models, rule-based accounts assume that people store a mental representation of frequency data and can use this information flexibly to compute conditional probabilities or contingency indexes. Therefore, any evidence that participants exposed to a series of cause-effect pairings can later recall the frequencies of each of the trial types considered in Table 1 (*a*, *b*, *c*, and *d*) has usually been interpreted as supporting rule-based models over associative ones. Just as an example, although a classical study conducted by Price and Yates' (1995) concluded that associative elements were necessary to account for some types of contingency judgments, the authors nevertheless considered that frequency estimates begged for a completely different kind of explanation. In their own words, "[t]o answer the question 'How often have B and A co-occurred?' however, seems to suggest, if not require, a much different strategy for generating a judgment. Specifically, it suggests that one directly accesses co-occurrences of B and A, perhaps by

processes similar to those described in current exemplar models of memory" (Price & Yates, 1995, p. 1646).

Second, participants are able to assess not only causal judgments but also flexibly compute other indexes of the relationship between the cause and the effect such as, for example, the conditional probability of the effect given the cause (Gredebäck, Winman, & Juslin, 2000; Vadillo & Matute, 2007; Vadillo, Miller, & Matute, 2005). This implies that participants have encoded the information in a format upon which several statistical rules can be applied. At first glance, this flexibility fits better with the cognitive architecture proposed by rule-based models than with the more automatic and simple encoding and retrieval processes assumed to be at work in associative mechanisms.

Third, some studies have shown that manipulations that are known to have a strong impact on causal judgments sometimes have little or no effect on other types of judgments, such as conditional probability estimates, effect predictions in the presence of the cause, and, most importantly, participants' estimations of the frequency of each type of trial across the sequence of trials. For instance, it has been found that cue competition (i.e., the fact that increasing the contingency between a cause and an effect has a detrimental effect in the assessment of the causal role of other potential causes of the same effect), which can be readily observed in causal judgments, has no impact on conditional probability estimates (Gredebäck et al., 2000; Matute, Arcediano, & Miller, 1996; but see Cobos, Caño, López, Luque, & Almaraz, 2000) and trial-type frequency estimates (Ramos-Álvarez & Catena, 2005). Similarly, the overall probabilities of the cause, p(C), and the effect, p(E), which are known to bias causal judgments, seem to have no impact on trial-by-trial effect predictions made by participants across the learning phase (Allan, Siegel, & Tangen, 2005; Perales, Catena, Shanks, & González, 2005; but see Vadillo, Musca, Blanco, & Matute, 2011). The trial-type frequency estimates also seem to be immune to

effect-density biases (Crump, Hannah, Allan, & Hord, 2007). In a similar vein, Cándido et al. (2006) showed that exposing participants to aversive stimuli that induce a negative mood, intermixed with the cue-outcome pairings, can have a biasing effect on causal judgments, without a parallel impact on frequency estimates.

The fact that some manipulations influence certain dependent variables but not others has been taken as a clue for discovering the cognitive priority of some processes over others. For example, in light of the fact that cue competition and emotion induction have an influence on causal judgments but not on frequency estimates, researchers have assumed that causal judgments are based on frequency estimates (and not the opposite) and whatever mechanisms are responsible for these effects, they must be acting at a relatively late stage of processing, before the causal judgment is produced but after the frequency data have been properly encoded and stored (Cándido et al., 2006, Ramos-Álvarez & Catena, 2005). Following the same logic, other researchers have concluded that estimations of the probability of the effect given the cause and effect predictions based on the presence of the cause must be based on earlier (and more automatic) cognitive processes than causal judgments (Allan et al., 2005; Gredebäck et al., 2000; Perales et al., 2005).

This interpretation fits nicely with the cognitive architecture underlying rule-based models of causal judgment: Participants would first store a representation of the raw frequency data and then they would use this information to compute conditional probabilities and contingency indexes. Manipulations such as cue interaction or emotion induction would influence causal judgments at this latter point. In contrast, it is difficult for associative models to explain why there are manipulations that only affect some judgments but not others. If participants' judgments are based on associations, shouldn't all types of judgments be biased in a similar manner? Moreover, why should participants have memories of frequency data at all?

Inferring frequency data from probabilities and associations

The preceding evidence has been generally taken in support for the idea that people must store some kind of mental representation of the raw data contained in a contingency matrix and that this information is later used to make inferences about probabilities or contingencies. However, as we will try to show in the present section, the opposite view is also plausible and cannot be disregarded in light of just this evidence. The fact that participants remember the frequencies of each trial type and that they are able to judge probabilities and contingencies does not, per se, prove that the former are the basis for the latter. In fact, it could happen that during training participants encode some information about the different conditional probabilities that relate cause and effect and that, if asked to do so, they use this information to infer the frequencies that must have been experienced. In other words, maybe it is frequency estimates, and not causal judgments or conditional probability ratings what requires an inference; and, accordingly, maybe it is conditional probabilities or even causal relationships, and not frequency information, what are directly encoded in memory.

Consider the following situation. As most participants in a causal learning experiment, imagine that you are exposed to a series of trials in which a potential cause might be present or absent (e.g., a patient taking or not taking a medicine) and an effect is also present or absent (e.g., the patient suffers an allergic reaction or not). Based on the instructions given to you and on the general structure of the task, you suspect that your goal is to learn to use the information about the cause in order to better predict the effect. However, at the end of training, you are suddenly and unexpectedly asked to estimate the number of times you saw the cause and the effect together, the number of times you saw the cause but not the effect, the number of times you saw the effect without the cause, and the number of times both elements were absent.

Even if you did not pay attention to this information, it might nevertheless be relatively easy to make a good guess based on just a little information you do remember. For instance, you might remember that you were exposed to approximately 50 trials[2] and that about half of them were trials in which the patient took the medicine. Thus, you already know that you experienced about 25 medicine-present trials and 25 medicine-absent trials. Additionally, you might remember that the chances of suffering an allergic reaction were noticeably higher for patients taking the medicine, compared to those not taking it, although the probability of suffering the allergy was positive even for the patients that didn't take the medicine. Let's say that your guess is that 80% of the patients who took the medicine suffered the allergic reaction, but only 20% of those who didn't suffered it. Based on all this information, you could easily infer that about 20 of the 25 medicine-present trials must have been medicine-allergy pairings, while only 5 of the 25 medicine-absent trials must have been no medicine-allergy pairings.

Therefore, a good estimate of the relative frequencies of each type of trial can be made on the basis of some knowledge of a) the probability of the cause, $p(C)$, b) the probability of the effect given the cause, $p(E|C)$, and c) the probability of the effect given the absence of the cause, $p(E|\sim C)$. As shown by the previous example, the estimated relative frequency of type a trials can be computed as:

$$erf_a = p(C) \cdot p(E|C) \qquad \text{(Equation 3)}$$

The relative frequencies of other trial types can be computed in a similar fashion:

$$erf_b = p(C) \cdot [1 - p(E|C)] \qquad \text{(Equation 4)}$$

$$erf_c = [1 - p(C)] \cdot p(E|\sim C) \qquad \text{(Equation 5)}$$

$$erf_d = [1 - p(C)] \cdot [1 - p(E|\sim C)] \qquad \text{(Equation 6)}$$

If you also have a general feeling of the total amount of trials you may have seen, it is very easy to infer even the absolute frequencies of each of these trial types, simply by multiplying this total amount of trials by each relative frequency.

As we have already discussed, the asymptotic value of the cause-effect association converges to the value of the objective cause-effect contingency. In a similar vein, other probabilistic parameters that relate the cause and the effect, apart from contingency, can be estimated on the basis of associations computed by means of the RW rule. For instance the probability of the effect given the cause, $p(E|C)$, can be computed adding the associative strength of the cause and the strength of the association between the context and the effect:

$$p(E|C) \approx V_{C\text{-}E} + V_{CTX\text{-}E} \qquad \text{(Equation 7)}$$

This is a natural consequence of the fact that in cause-present trials the RW rule tries to minimize the error made when predicting the effect based on the associative strengths of the cause and the context. This can only be accomplished by gradually developing cause-effect and context-effect associative strengths whose sum is $p(E|C)$.[3]

Similarly, in cause absent trials, the RW rule minimizes the error made in predicting the effect on the basis of the context-effect association, which can only be accomplished by developing a context-effect association whose associative strength approaches $p(E|\sim C)$. Therefore, the probability of the effect in the absence of the cause is asymptotically equivalent to the associative strength of the context,

$$p(E|\sim C) \approx V_{CTX\text{-}E} \qquad \text{(Equation 8)}$$

For the same reason, if there is a single, constant context, the overall probability of the cause, $p(C)$, is equal to the probability of the cause given the context, $p(C|CTX)$, which is equivalent to the asymptotic value of the association between the context and the cause,

$$p(C) \approx V_{CTX\text{-}C} \qquad \text{(Equation 9)}$$

As in Equations 7 and 8, Equation 9 follows from the fact that the context-cause association computed with the RW rule would minimize the error made when trying to predict the cause on the basis of the context. This error is minimal when the associative strength of the context-cause association equals the probability of the cause given the context, $p(C|CTX)$, which, in situations in which the context is constant also equals the overall probability of the cause, $p(C)$. The strength of this association, $V_{CTX\text{-}C}$, can also be computed with the general RW rule using λ to code for the presence or absence of the cause (instead of the effect), β to represent the salience of the cause (also instead of that of the effect) and V_{TOTAL} to represent the strength of the context-cause association in the previous trial.

Equations 7-9 can be combined with Equations 3-6, so that the estimated relative frequencies of each trial type can be inferred from the value of these three associations,

$$erf_a \approx V_{CTX\text{-}C} \cdot (V_{C\text{-}E} + V_{CTX\text{-}E}) \qquad \text{(Equation 10)}$$

$$erf_b \approx V_{CTX\text{-}C} \cdot [1 - (V_{C\text{-}E} + V_{CTX\text{-}E})] \qquad \text{(Equation 11)}$$

$$erf_c \approx (1 - V_{CTX\text{-}C}) \cdot (V_{CTX\text{-}E}) \qquad \text{(Equation 12)}$$

$$erf_d \approx (1 - V_{CTX\text{-}C}) \cdot (1 - V_{CTX\text{-}E}) \qquad \text{(Equation 13)}$$

which means that, contrary to the logic followed by the authors of the studies previously considered, the fact that people are able to "recall" the frequency of each trial type need not be taken as evidence that these data are directly encoded in their memory and that, therefore, all other estimates have to be inferences made on the basis of this information. At least from a formal point of view, the possibility that causal judgments are based on encoded frequency estimates is just as likely as the opposite: That frequency estimates are the result of an inference made on the basis of probability information, which could be encoded as associations formed by means of the RW rule. Equations 10-13 are just as an example of how this can be done. The following simulations were conducted in order to

better assess the ability of Equations 10-13 to correctly infer frequency estimates under a number of conditions.

Simulation 1: Frequency assessment under different contingencies

First, we were interested in knowing how these frequency-retrieval rules behaved under different cause-effect contingencies. Therefore, we conducted a simulation comprising four conditions in which the probability of the effect given the cause and the probability of the effect given the absence of the cause were manipulated, yielding four different contingency values.

Method

The design summary of the conditions included in Simulation 1 is shown in Table 2. As can be seen, the simulation included four conditions. Each condition consisted of a different combination of frequencies of each type of trial (*a*, *b*, *c*, and *d*) that resulted in different values for $p(E|C)$ and $p(E|\sim C)$ and, consequently, in a different cause-effect contingency, as defined by Δp. The two numbers used to denote the experimental conditions refer to the probability of the effect given the cause, $p(E|C)$ and the probability of the effect in the absence of the cause, $p(E|\sim C)$, respectively. In two conditions (.20-.50 and .20-.80), the cause-effect contingency was negative (-.30 and -.60, respectively) and in the other two (.50-.20 and .80-.20) the contingency was positive (.30 and .60, respectively). Each condition comprised a sequence of 100 trials.

In most simulations of the RW rule, the learning rate parameter α is larger for the cause than for the constant context, whose salience is assumed to be relatively low (e.g., Mercier, 1996). Therefore, the learning-rate parameter α was set to 0.5 for the cause and 0.2 for the context. The learning-rate parameter β was set to 0.5 for the computation of all the associations (cause-effect, context-effect, and context-cause associations).

In order to properly measure the models ability to infer the relative frequencies of each trial type, in each trial we computed the root mean squared error made by the model when estimating the relative frequency of trials *a*-*d* on the basis of Equations 10-13. This error was computed according to the following equation:

$$Error = \sqrt{\frac{\sum_{i=a,b,c,d}(erf_i - rf_i)^2}{4}} \qquad \text{(Equation 14)}$$

That is, the error was computed as the squared root of the mean squared difference between the estimation of relative frequencies, erf_i, and actual relative frequencies, rf_i.

On each simulated trial, the program first chose a trial type (*a*, *b*, *c*, or *d*) from the list of trials in that condition (see Table 2). Then, the strengths of the cause-effect association, the context-effect association, and the context-cause associations were updated following Equation 2 and using the learning-rate parameters mentioned above. After updating the associative strengths, the frequency-retrieval rules instantiated in Equations 10-13 were used to reconstruct the relative frequencies of each trial type that could be inferred on the basis of those associations. Finally, the error in these inferences was computed following Equation 14, and the simulation proceeded to the next simulated trial. In order to obtain smooth learning curves, we conducted five hundred simulations for each condition, each one with a randomly ordered sequence of trials.

Results and Discussion

The results of the simulation are depicted in Figure 1. The top panel shows the strengthening of the cue-outcome association in each condition over trials. As could be expected, this associative strength gradually approaches the value of the objective cue-outcome contingency, as predicted by the computational analysis of the Rescorla-Wagner model mentioned in the Introduction (Chapman & Robbins, 1990; Danks, 2003; Wasserman et al., 1993).

The lower panel shows the average error made by the model when estimating the frequency of each trial type. As can be seen, this error is large in the beginning of each simulation, before the associative strengths used in the estimations have reached their asymptotic value. However, as soon as the three associative strengths, $V_{C\text{-}E}$, $V_{CTX\text{-}E}$, and $V_{CTX\text{-}C}$, reach the learning asymptote, the global error made in the frequency assessment decreases noticeably in all conditions. Interestingly, the global errors tend to be rather similar in both conditions with positive contingency, but they are systematically larger in negative contingencies. In fact, the more negative the contingency, the larger this overall error seems to be. Figure 2 depicts the estimated relative frequencies of each trial type in Simulation 2.

Simulation 2: Frequency assessment and outcome-density effects

Most of the evidence arguing that people store a memory trace of trial frequencies is based on manipulations that affected causal judgment but leave frequency assessments unaltered (Cándido et al., 2006; Price & Yates, 1995; Crump et al., 2007; Ramos-Álvarez & Catena, 2005). In our next simulation we wanted to show that this type of evidence is not at odds with the Rescorla-Wagner learning algorithm *per se.* Specifically, we wanted to assess to what extent a well-studied manipulation that is known to have a strong impact in causal judgment, namely the outcome-density effect (e.g., Alloy & Abramson, 1979; Allan & Jenkins, 1983; Matute, Yarritu, & Vadillo, 2011; Musca, Vadillo, Blanco, & Matute, 2010; Wasserman, Kao, van Hamme, Katagiri, & Young, 1996), does also have an influence on frequency estimates. Therefore, we contrasted three conditions, all of them with the same cause-effect contingency, $\Delta p = 0.50$, but each one with a different overall probability of the effect. It is widely-known that the Rescorla-Wagner learning algorithm predicts that this manipulation should have at least a preasymptotic biasing effect on the development of cue-outcome associations. However, as we will try to show in this

simulation, this need not imply that the same biasing effect should appear in frequency estimations made on the basis of these associations.

Method

The specific probabilities of the effect, in the presence and in the absence of the cause, in each of the three conditions included in Simulation 2 are shown in Table 2, along with the frequencies of each trial type, *a*-*d*. In this case, each condition comprised a sequence of 80 trials, instead of 100. All the other procedural details (the learning-rate parameters and the number of simulations per condition) were kept the same as in Simulation 1.

Results and Discussion

As can be seen in the top panel of Figure 3, the strength of the cue-outcome association is noticeably biased by this manipulation, though only preasymptotically. This is a well-known property of the RW model (Shanks, 1995). However, as shown in the lower panel, this outcome-density manipulation has little impact on frequency estimations. Furthermore, the condition that should give rise to larger outcome-density biases, 1.00-0.50, is actually the one with the best overall performance; and there are few, if any, differences between the other two conditions. This happens because, even though cue-outcome associations develop at a different rate in each condition, the frequency estimates do not only depend on that single association, but in a complex equilibrium of associations that can remain globally accurate in spite of minor individual preasymptotic biases in each association.

This last simulation is particularly relevant because it shows that a manipulation can have a biasing effect on a cue-outcome association without a parallel effect on other inferences that can be made on the basis of that association (in combination with other associations). Thus, all the evidence presented above showing that there are manipulations

that have an effect on contingency or causal judgments without a similar impact on frequency estimates (Cándido et al., 2006; Crump et al., 2007; Price & Yates, 1995; Ramos-Álvarez & Catena, 2005) or other dependent variables (Allan et al., 2005; Gredebäck et al., 2000; Perales et al., 2005; Vadillo & Matute, 2007; Vadillo et al., 2005), need not imply that these latter judgments are cognitively simpler than the former. Nor do they imply that causal judgments are based on a memory trace of the frequency data, instead of the opposite.

Simulation 3: Manipulating the learning rates of cues and contexts

One of the experiments that we have discussed above (Cándido et al., 2006) found that the presentation of aversive stimuli intermixed with the cue-outcome pairings delayed the perception of contingency, but did not have a parallel effect on frequency estimations. As the authors argue, associative models could potentially account for the former effect by means of a change in the learning rates. For example, in the RW model the learning rate depends on the salience of the cause (α_{cue}), the salience of the outcome (β), and the salience of a constant context ($\alpha_{context}$). However, according to Cándido et al. (2006), as the participants also remembered the frequency information and, moreover, this memory was not affected by the aversive stimulation, these data beg for a nonassociative explanation. Contrary to this claim, in Simulation 3 we show that even from our associative view, manipulating the learning rates can have complex effects on the cause-effect association and on the frequency estimations. While some manipulations do have a similar effect in both variables, some other manipulations have a stronger impact on the cause-effect association than in the frequency estimations, consistent with the results of Cándido et al. (2006).

Method

In all the conditions, the contingency was kept constant at 0.60, with the following trial frequencies: 40 *a* trials, 10 *b* trials, 10 *c* trials, and 40 *d* trials. As in Simulations 1 and 2, the learning-rate parameter β was set to 0.5 and 500 randomly-ordered trial sequences were simulated for each condition. The saliences of the cause and of the context were manipulated orthogonally. α_{cause} was 0.2 for half of the conditions and 0.8 for the other half. Similarly, $\alpha_{context}$ was 0.2 for half of the conditions and 0.8 for the other half.

Results and Discussion

Figure 4 shows the results of Simulation 3. The first number in the name of the data series refers to the salience of the cause (α_{cause} = 0.2 or 0.8) and the second number to the salience of the context ($\alpha_{context}$ = 0.2 or 0.8). As can be seen in the top panel, the manipulation of the learning rates had a clear impact on the cause-effect association. The conditions with the higher α_{cause} were the first ones to reach the learning asymptote. The salience of the context, also had an impact (though seemingly smaller) in the acquisition of the cause-effect association.

The lower panel shows that the manipulation of the learning rates also had a remarkable effect on the relative frequency estimations, although the effects on this variable are somewhat more complex. The condition in which both learning rates were highest (.8/.8) yields the worse frequency estimations. The reason for this is that the high learning rates result in very large oscillations around the learning asymptote after every trial (recall the smooth learning curves that can be seen in the top panel of Figure 4 show the average results of 500 simulations with different trial orders), which in turn gives rise to less than perfect frequency estimations on each trial. By contrast, in the condition with the lowest learning rates, .2/.2, these oscillations are minimal and have a less disturbing effect on the relative frequency estimations.

The most interesting conditions for our present purposes are those in which the cause and the context have different saliences. As can be seen in the top panel, conditions .8/.2 and .2/.8 have the most different cause-effect associations; however, they show a relatively similar accuracy in their relative frequency estimations. This shows that manipulations of the learning rates can have very different effects on the development of the cause-effect association and on the frequency estimates: Manipulations that have the stronger impact on the former might have a rather small impact on the latter, and viceversa. Thus, contrary to the theoretical interpretation of Cándido et al. (2006), their dissociation between contingency judgments and frequency estimations, can be accounted for in purely associative terms.

Simulation 4: Frequency estimations based on other associative-learning rules

The previous simulations rely in the RW learning rule, which is perhaps the most popular associative learning theory in the area of animal conditioning and also in human contingency learning research (Allan, 1993; Shanks, López, Darby, & Dickinson, 1996). However, the same idea can be implemented in alternative learning algorithms. For example, the comparator hypothesis (Miller & Matzel, 1988; Stout & Miller, 2007) assumes that cause-effect associations are developed by means of a much simpler learning algorithm which is only sensitive to the conditional probability of the outcome given the cue, instead of the cause-effect contingency (Bush & Mosteller, 1951). It is easy to develop an associative learning model that reconstructs the frequency data on the basis of cause-effect associations computed with that algorithm.

Similarly, some researchers (Cheng, 1997; Holyoak & Cheng, 2011; Novick & Cheng, 2004) have argued that human causal learning is not sensitive to the cause-effect contingency, but to a related concept called causal power. Under some conditions, the

generative causal power (i.e., the ability of a cause to produce an effect) can be computed from the observable contingency information by means of the following rule:

$$q = \frac{\Delta P}{1 - p(E \mid \sim \mathrm{C})} \qquad \text{(Equation 15)}$$

In principle, this equation is not proposed as a specific rule that people apply consciously to infer causal relations, but as a normative standard for the rational analysis of causal learning. However, some authors have proposed an associative learning rule, similar to the RW model, that computes causal power asymptotically and that could provide an algorithmic explanation of peoples' ability to estimate causal power (Danks, Griffiths, & Tenenbaum, 2003). As we try to show below, it is equally plausible to make inferences about the relative frequencies of each trial type on the basis of the output of this associative learning rule.

The associative rule proposed by Danks et al. (2003) only differs from the RW rule in the assumptions about how the associative strengths of several stimuli (e.g., the cause and the context) are combined to produce V_{TOTAL}. While in the standard RW model V_{TOTAL} is linear sum of the associative strengths of the stimuli that are present on any single trial, in the model proposed by Danks et al. (2003) these associative strengths are combined by means of a noisy-OR integration rule.[4] In situations in which only one potential cause and a constant context are involved, this rule equals to computing V_{TOTAL} as:

$$V_{TOTAL} = V_{C\text{-}E} + V_{CXT\text{-}E} - (V_{C\text{-}E} \cdot V_{CXT\text{-}E}) \qquad \text{(Equation 16)}$$

When this associative learning rule is used instead of the standard RW rule, the probability of the effect given the cue, $p(E|C)$, can no longer be estimated following Equation 7, but it can be computed as follows:

$$p(E|C) \approx V_{C\text{-}E} + V_{CTX\text{-}E} - (V_{C\text{-}E} \cdot V_{CXT\text{-}E}) \qquad \text{(Equation 17)}$$

Consequently, Equations 10 and 11 also need to be rewritten as:

$$erf_a \approx V_{CTX\text{-}C} \cdot [V_{C\text{-}E} + V_{CTX\text{-}E} - (V_{C\text{-}E} \cdot V_{CXT\text{-}E})] \qquad \text{(Equation 18)}$$

$$erf_b \approx V_{CTX\text{-}C} \cdot \{1 - [V_{C\text{-}E} + V_{CTX\text{-}E} - (V_{C\text{-}E} \cdot V_{CXT\text{-}E})]\} \qquad \text{(Equation 19)}$$

In this simulation we show that, as expected, when using these equations to compute the strength of the cause-effect association, its asymptotic value depends on causal power (as computed by Equation 15) and not on contingency. Furthermore, regardless of the precise value of causal power or contingency, the frequency estimations that can be made on the basis of this associative learning algorithm (based on Equations 12, 13, 18, and 19) are similarly accurate across conditions and do not differ remarkably from those observed in Simulations 1-3, when the standard RW rule was used.

Method

Four conditions were included in Simulation 4. In half of the conditions, the generative power of the cause, as computed by Equation 15, was 0.25 and in the other half it was 0.50. Ortogonally, in half of the conditions the cause-effect contingency as measured by Δp was 0.25 and in the other half it was 0.10. The trial frequencies of each condition are shown in Table 1. All the parameter values were kept as in Simulations 1 and 2. As in previous simulations, 500 replications, each one with a randomized trial sequence, were simulated for each condition.

Results and Discussion

The results of Simulation 4 are depicted in Figure 5. The first number in the name of each data series refers to the causal power and the second one to the Δp value of that condition. As can be seen in the top panel of Figure 5, the asymptotic associative strength converges to the causal power, with contingency playing a relatively minor role in the preasympotic strength of the associations. The lower panel shows that the accuracy of the relative frequency estimations made on the basis of the associations was relatively similar in all conditions and rather similar to those found in previous associations when the standard RW rule was used for the computation of the associative strengths. Interestingly,

conditions with remarkably different learning curves, such as, for example, conditions .50/.25 and .25/.10 reach similar levels of accuracy in the relative frequency estimations. This provides additional support for our claim that manipulations that influence the value of the cause-effect association need not have a similar impact on the accuracy of frequency estimates.

General Discussion

In the Introduction, we argued that researchers have often made use of three arguments to favor rule-based accounts of causal learning over associative explanations. First, the fact that participants can estimate the number of trials of each type they have been exposed to is assumed to fit better with rule based accounts. Admittedly, the associative models most widely cited in the area of causal learning do not include any mechanism by which frequency data can be encoded and stored. However, as we have shown, the associations computed by means of the Rescorla-Wagner rule and related associative algorithms contain all the information necessary to the reconstruct the frequency data. Therefore, any evidence that people "remember" this information cannot be taken as a strong argument in favor of rule-based models.

Second, the abundant literature showing that people can flexibly compute several statistical indexes to describe the relationship between the cause and the effect (Gredebäck et al., 2000; Matute et al., 1996; Vadillo et al., 2005; Vadillo & Matute, 2007) has sometimes been interpreted as uniquely supporting rule-based accounts over associative ones (see, e.g., Cobos et al., 2000; Gredebäck et al., 2000). However, just the same way associations contain all the information necessary to reconstruct the contingency table that gave rise to those associations, that information can also be used to infer other statistical indexes relating cause and effect. For example, as noted by Vadillo and Matute (2007), the probability of the effect given the cause can be computed by combining the associative

strengths of the cause and the context. Moreover, this associative perspective makes new predictions that cannot be easily accommodated by rule-based models without making *ad hoc* assumptions. The associative framework described in the present paper is nothing but an extension of these ideas.

Perhaps the strongest arguments in favor of rule-based accounts are the ones related to the third point mentioned in the introduction: That is, results showing that there are experimental manipulations that have an effect on some dependent variables but not others (Gredebäck et al., 2000; Matute et al., 1996; Pineño, Denniston, Beckers, Matute, & Miller, 2005). Dissociations between causal judgments and frequency estimations are particularly compelling for the present work (Cándido et al., 2006; Crump et al., 2007; Price & Yates, 1995; Ramos-Álvarez & Catena, 2005). However, our Simulation 2, aimed at addressing this particular issue, illustrates how, at least sometimes, inferences made on the basis of biased associations need not exhibit the same biases. Specifically, we showed that, although an outcome-density manipulation does have an effect on the cause-effect associations computed by the Rescorla-Wagner algorithm, that manipulation has no observable impact on the frequency estimations based on those associations. Similarly, in Simulation 3 we showed that manipulations of the learning rates need not have the same impact on cause-effect associations and on the relative frequency estimations. The results of Simulation 4 also support this view.

Some of the published reports that have used dissociations in judgments to test associative models have relied on cue competition manipulations. In a typical cue competition design, several cues (or potential causes) are trained as predictors of the outcome (or effect), with the usual result that the contingency between one of those cues with the outcome has an effect not only on judgments about that outcome, but also on judgments about the second outcome, whose contingency is not manipulated. As

mentioned above, extant evidence shows that although cue competition affects causal and contingency learning, again it has little or no impact on predictive judgments (Gredebäck et al., 2000; Matute et al., 1996; Pineño, Denniston, Beckers, Matute, & Miller, 2005) or on frequency estimates (Ramos-Álvarez & Catena, 2005). The frequency-retrieval rules that we have developed here deal only with situations involving just one cue and one context and, therefore, they are not well suited to account for these dissociations. However, as the astute reader may have noticed, the logic behind this set of rules can be easily extended to multiple-cue settings, so that unbiased estimations of frequency data can be observed in multiple cause settings that usually give rise to cue competition effects. In this case, frequency estimates would rely not only in three associations but more, including associations of the context with each cue, associations between both cues, and so on. Such a model would certainly look more complex from a purely formal point of view, but would be based on a similarly simple logic: Instead of recalling frequency data, they would just infer these data from their intuitions about conditional probabilities.

Regardless of the potential merits and shortcomings of the specific formalization that we have advanced here, our simulations show that researchers should be more cautious when making strong claims about incompatibility of associative models with any evidence of flexibility of judgments. If we want to have a more detailed picture of the cognitive priority of some processes over others, then we cannot simply rely on information about these dissociations. This strategy should be complemented with alternative methods that allow us to more directly assess the relative complexity or automaticity of several processes and their precise sequence over time. For example, an important (although not necessary) assumption behind associative models is that learning mechanisms operate in a relatively automatic fashion and posse few demands in terms of cognitive resources. Based on this idea many researchers have tested the plausibility of these mechanisms by trying to

measure learning effects using more implicit measures, instead of the traditional verbal causal ratings. Some of these studies (e.g., Morís, Cobos, & Luque, 2010; Sternberg & McClelland, 2012) have obtained data consistent with associative models, while others have failed to find some learning effects with these implicit measures (e.g., De Houwer & Vandorpe, 2010; Ratliff & Nosek, 2010). Although the evidence is not conclusive yet, we think that these alternative measures of learning are a promising tool for any attempt to discriminate between associative and rule-based accounts of contingency learning. Manipulations of time-pressure (Vadillo & Matute, 2010) or secondary task (De Houwer & Beckers, 2003; Wills, Graham, Koh, McLaren, & Rolland, 2011) are complementary tools to address the automaticity of the processes involved in contingency learning. Similarly, physiological data can also be used to test the plausibility of different theories by looking for physiological correlates of the processes assumed by each theory. For example, recent research has found correlates of error-correction processes or attention modulation that are consistent with associative learning theory (e.g., Fletcher et al., 2001; Luque, López, Marco-Pallares, Càmara, & Rodríguez-Fornells, in press; Walsh & Anderson, 2011; Wills, Lavric, Croft, & Hodgson, 2007). We think that any serious attempt to contrast the predictions of rule-based and associative accounts of contingency learning should benefit from the fruitful combination of these and other methods, instead of relying solely in relatively simplistic interpretations of judgment dissociations which, as we have tried to show in the present paper, are usually open to alternative and equally plausible interpretations.

References

Allan, L.G. (1980). A note on measurement of contingency between two binary variables in judgement tasks. *Bulletin of the Psychonomic Society, 15,* 147-149.

Allan, L. G. (1993). Human contingency judgments: Rule-based or associative? *Psychological Bulletin, 114,* 435-448.

Allan, L. G. (2003). Assessing power PC. *Learning & Behavior, 31,* 192-204.

Allan, L. G., & Jenkins, H. M. (1980). The judgment of contingency and the nature of the response alternatives. *Canadian Journal of Psychology, 34,* 1-11.

Allan, L. G., & Jenkins, H. M. (1983). The effect of representations of binary variables on judgment of influence. *Learning and Motivation, 14,* 381-405.

Allan, L. G., Siegel, S., & Tangen, J. M. (2005). A signal detection analysis of contingency data. *Learning & Behavior, 33,* 250-263.

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.

Beyth-Marom, R. (1982). Perception of correlation re-examined. *Memory & Cognition, 10,* 511-519.

Busemeyer, J. R. (1991). Intuitive statistical estimation. In N. H. Anderson (Ed.), Contributions to information integration theory: Vol. 1. Cognition (pp. 187-215). Hillsdale, NJ: Erlbaum.

Cándido, A., Perales, J. C., Catena, A., Maldonado, A., Guadarrama, L., Beltrán, R., Contreras, D., & Herrera, A. (2006). Efectos de inducción emocional en el aprendizaje causal [Effects of emotion induction on causal learning]. *Psicológica, 27,* 243-267.

Chapman, G. B., & Robbins, S. J. (1990). Cue interaction in human contingency judgment. *Memory & Cognition, 18,* 537-545.

Cheng, P. W. (1997). From covariation to causation: a causal power theory. *Psychological Review, 104,* 367-405.

Cheng, P. W., & Novick, L. R. (1992). Covariation in natural causal induction. *Psychological Review, 99,* 365-382.

Cobos, P. L., Caño, A., López, F. J., Luque, J. L., & Almaraz, J. (2000). Does the type of judgement required modulate cue competition? Quarterly Journal of Experimental Psychology, 53B, 193-207.

Crump, M. J. C., Hannah, S. D., Allan, L. G., & Hord, L. K. (2007). Contingency judgments on the fly. *Quarterly Journal of Experimental Psychology, 60,* 753-761.

Danks, D. (2003). Equilibria of the Rescorla-Wagner model. *Journal of Mathematical Psychology, 47,* 109-121.

Danks, D., Griffiths, T. L., & Tenenbaum, J. B. (2003). Dynamical causal learning. In S. Becker, S. Thrun, & K. Obermayer (Eds.), *Advances in neural information processing systems* (Vol. 15, pp. 67-74). Cambridge, MA: MIT Press.

De Houwer, J., & Beckers, T. (2003). Secondary task difficulty modulates forward blocking in human contingency learning. *Quarterly Journal of Experimental Psychology, 56B*, 345-357.

De Houwer, J., & Vandorpe, S. (2010). Using the Implicit Association Test as a measure of causal learning does not eliminate effects of rule learning. *Experimental Psychology, 57,* 61-67.

Dickinson, A., Shanks, D. R., & Evenden, J. (1984). Judgement of act-outcome contingency: The role of selective attribution. *Quarterly Journal of Experimental Psychology, 36A,* 29-50.

Fletcher, P. C., Anderson, J. M., Shanks, D. R., Honey, R., Carpenter, T. A., Donovan, T., Papadakis, N., & Bullmore, E. T. (2001). Responses of human frontal cortex to surprising events are predicted by formal associative learning theory. *Nature Neuroscience, 4,* 1043–1048.

Gredebäck, G., Winman, A., & Juslin, P. (2000). Rational assessments of covariation and causality. *Proceedings of the 22nd Annual Conference of the Cognitive Science Society* (pp. 190-195). Hillsdale, NJ: Lawrence Erlbaum Associates.

Holyoak, K. J., & Cheng, P. W. (2011). Causal learning and inference as a rational process: The new synthesis. *Annual Review of Psychology, 62,* 135-163.

Hume, D. (1964). *Treatise of human nature* (ed., L. S. Selby-Bigge). London: Oxford University Press. (Original work published 1739.)

Jenkins, H. M., & Ward, W. C. (1965). Judgment of contingency between responses and outcomes. *Psychological Monographs, 79,* 1- 17.

Luque, D., López, F. J., Marco-Pallares, J., Càmara, E., & Rodríguez-Fornells, A. (in press). Feedback-related brain potential activity complies with basic assumptions of associative learning theory. *Journal of Cognitive Neuroscience.*

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information.* Freeman, San Francisco.

Matute, H., Arcediano, F., & Miller, R. R. (1996). Test question modulates cue competition between causes and between effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 182-196.

Matute, H., Yarritu, I., & Vadillo, M. A. (2011). Illusions of causality at the heart of pseudoscience. *British Journal of Psychology, 102,* 392-405.

Mercier, P. (1996). Computer simulations of the Rescorla-Wagner and Pearce-Hall models in conditioning and contingency judgment. *Behavior Research Methods*, 28, 55-60.

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, CA: Academic Press.

Morís, J., Cobos, P. L., & Luque, D. (2010). What priming techniques can tell us about associative representations acquired during human contingency learning. *Open Psychology Journal, 3,* 97-104.

Musca, S. C., Vadillo, M. A., Blanco, F., & Matute, H. (2010). The role of cue information in the outcome-density effect: Evidence from neural network simulations and a causal learning experiment. *Connection Science, 22*, 177-192.

Novick, L. R., & Cheng, P. W. (2004). Assessing interactive causal influence. *Psychological Review, 111,* 455-485.

Pearl, J. (2000). *Causality: Models, reasoning, and inference.* New York: Cambridge University Press.

Perales, J. C., Catena, A., Shanks, D. R., & González, J. A. (2005). Dissociation between judgments and outcome expectancy measures in covariation learning: A signal detection theory approach. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 1105-1120.

Pineño, O., Denniston, J. C., Beckers, T., Matute, H., & Miller, R. R. (2005). Contrasting predictive and causal values of predictors and causes. *Learning & Behavior, 33,* 184-196.

Price, P. C., & Yates, J. F. (1995). Associative and rule-based accounts of cue interaction in contingency judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 1639-1655.

Ramos-Álvarez, M. M., & Catena, A. (2005). The dissociation between the recall of stimulus frequencies and the judgment of contingency allows the placement of the competition effect in the final causal processing stages. *Psicológica, 26,* 293-303.

Ratliff, K. A., & Nosek, B. A. (2010). Creating distinct implicit and explicit attitudes with an illusory correlation paradigm. *Journal of Experimental Social Psychology, 46,* 721-728.

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.

Shaklee, H., & Mims, M. (1982). Sources of error in judging event covariations: Effects of memory demands. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8,* 208-224.

Shanks, D. R. (1987). Acquisition functions in contingency judgment. *Learning and Motivation, 18,* 147-166.

Shanks, D. R. (1995). Is human learning rational? *Quarterly Journal of Experimental Psychology, 48A,* 257-279.

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

Shanks, D. R., & Dickinson, A. (1991). Instrumental judgment and performance under variations in action-outcome contingency and contiguity. *Memory & Cognition, 19,* 353-360.

Shanks, D. R., López, F. J., Darby, R. J., & Dickinson, A. (1996). Distinguishing associative and probabilistic contrast theories of human contingency judgment. In D.

R. Shanks, Holyoak, K. J., & D. L. Medin (Eds.), *The psychology of learning and motivation, Vol. 34: Causal learning* (pp. 265-311). San Diego, CA: Academic Press.

Smedslund, J. (1963). The concept of correlation in adults. *Scandinavian Journal of Psychology, 4,* 165-173.

Sternberg, D. A. & McClelland, J. L. (2012). Two mechanisms of human contingency learning. *Psychological Science,* 23, 59-68.

Stout, S. C., & Miller, R. R. (2007). Sometimes-competing retrieval (SOCR): A formalization of the comparator hypothesis. *Psychological Review, 114,* 759-783.

Vadillo, M. A., & Matute, H. (2007). Predictions and causal estimations are not supported by the same associative structure. *Quarterly Journal of Experimental Psychology, 60,* 433-447.

Vadillo, M. A., Miller, R. R., & Matute, H. (2005). Causal and predictive-value judgments, but not predictions, are based on cue-outcome contingency. *Learning & Behavior, 33,* 172-183.

Vadillo, M. A., Musca, S. C., Blanco, F., & Matute, H. (2011). Contrasting cue-density effects in causal and prediction judgments. *Psychonomic Bulletin & Review, 18,* 110-115.

Walsh, M. M., & Anderson, J. R. (2011). Modulation of the feedback-related negativity by instruction and experience. *Proceedings of the National Academy of Sciences, 108,* 19048-19053.

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: Academic Press.

Wasserman, E. A., Elek, S. M., Chatlosh, D. L., & Baker, A. G. (1993). Rating causal relations: The role of probability in judgments of response-outcome contingency.

Journal of Experimental Psychology: Learning, Memory, and Cognition, 19, 174-188.

Wasserman, E. A., Kao, S.-F., Van Hamme, L. J., Katagari, M., & Young, M. E. (1996). Causation and association. In D. R. Shanks, K. J. Holyoak, & D. L. Medin (Eds.), *The psychology of learning and motivation, Vol. 34: Causal learning* (pp. 207-264). San Diego, CA: Academic Press.

White, P. A. (2003). Making causal judgments from the proportion of confirming instances: The pCI rule. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 710-727.

Wills, A. J., Graham, S., Koh, Z., McLaren, I. P. L., & Rolland, M. D. (2011). Effects of concurrent load on feature- and rule-based generalization in human contingency learning. *Journal of Experimental Psychology: Animal Behavior Processes, 37,* 308-316.

Wills, A. J., Lavric, A., Croft, G. S., & Hodgson, T. L. (2007). Predictive learning, prediction errors, and attention: Evidence from event-related potentials and eye tracking. *Journal of Cognitive Neuroscience, 19,* 843-854.

Author Note

Support for this research was provided by Grant IT363-10 from Departamento de Educación, Universidades e Investigación of the Basque Government and Grants PSI2011-26965 and PSI2011-24662 from Ministerio de Ciencia e Innovación. We would like to Lorraine Allan and an anonymous reviewer for their helpful comments on a previous draft of this article. Correspondence concerning this article should be addressed to Miguel A. Vadillo, Departamento de Fundamentos y Métodos de la Psicología, Universidad de Deusto, Apartado 1, 48080 Bilbao, Spain. E-mail: mvadillo@deusto.es

Footnotes

[1] This prediction only holds if the β learning-rate parameter is assumed to have the same value in effect-present and effect-absent trials.

[2] In fact, this information is usually given to participants in many causal learning studies. For example, when participants are presented with information about a fictitious patient taking a medicine and suffering or not an allergic reaction, it is not uncommon that each trial begins with the sentence "On day X, the patient took the medicine". Therefore, participants can easily remember how many trials they have been exposed to by simply paying attention to the number of days.

[3] The simulation of the RW rule conducted by Vadillo and Matute (2007, p. 441) shows that conditions differing in their cause-effect contingencies but nevertheless similar regarding $p(E|C)$ give rise to different cause-effect associations, but similar values of the sum $V_{C\text{-}E} + V_{CTX\text{-}E}$.

[4] The formulation proposed by Danks et al. (2003) includes both a noisy-OR rule for the combination of generative causes and an AND-NOT integration rule for the combination of preventative causes. For the sake of simplicity, we have only simulated positive cause-effect relationships and, consequently, our formalization neglects the AND-NOT integration rule necessary to take into account preventive causal powers.

Table 1

Contingency table

	Effect present (E)	Effect absent (~E)
Cause present (C)	*a*	*b*
Cause absent (~C)	*c*	*d*

Table 2

Summary of the conditions included in Simulations 1, 2, and 4

Simulation 1

Condition	Frequencies	*p*(E|C)	*p*(E|~C)	Δ*p*
.20 - .50	10a, 40b, 25c, 25d	.20	.50	-.30
.20 - .80	10a, 40b, 40c, 10d	.20	.80	-.60
.50 - .20	25a, 25b, 10c, 40d	.50	.20	.30
.80 - .20	40a, 10b, 10c, 40d	.80	.20	.60

Simulation 2

Condition	Frequencies	*p*(E|C)	*p*(E|~C)	Δ*p*
.50 - .00	20a, 20b, 0c, 40d	.50	.00	0.50
.75 - .25	30a, 10b, 10c, 30d	.75	.25	0.50
1.00 - .50	40a, 0b, 20c, 20d	1.00	.50	0.50

Simulation 4

Condition	Frequencies	*p*(E|C)	*p*(E|~C)	*Q*	Δ*p*
.25/.25	25a, 75b, 0c, 100d	.25	.00	0.25	0.25
.25/.10	70a, 30b, 60c, 40d	.70	.60	0.25	0.10
.50/.25	75a, 25b, 50c, 50d	.75	.50	0.50	0.25
.50/.10	90a, 10b, 80c, 20d	.90	.80	0.50	0.10

Figure Captions

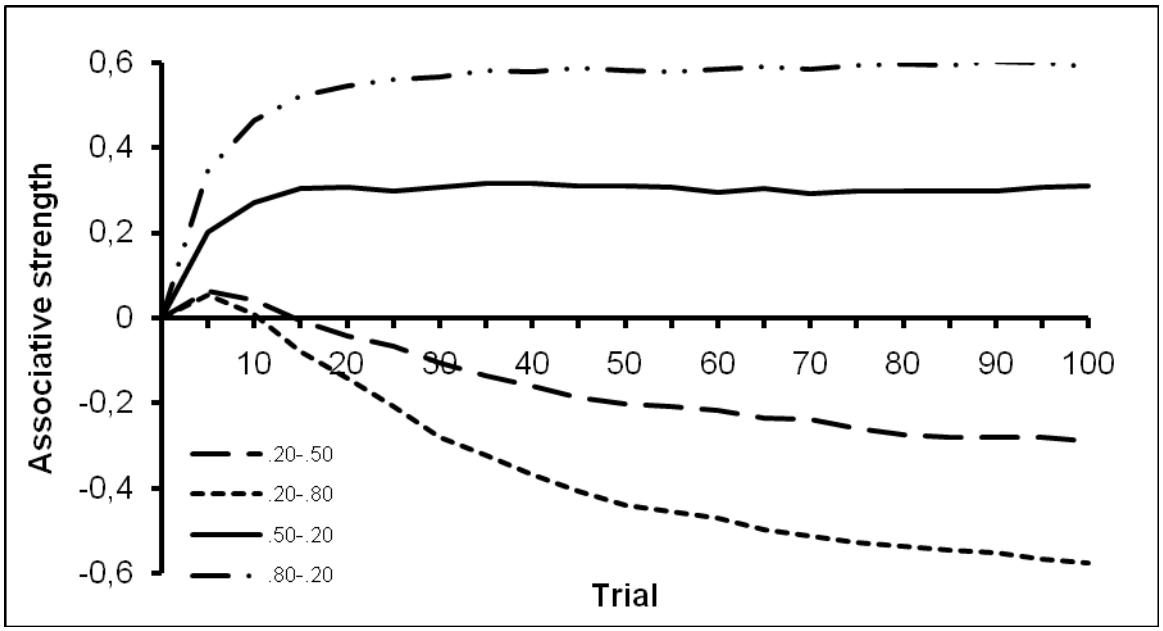
Figure 1. Results of Simulation 1. The top panel shows the strength of the cue-outcome association in each condition. The lower panel shows the global error made by the model when attempting to infer the relative frequency of each type of trial.

Figure 2. Results of Simulation 1. Estimated relative frequencies of each trial type (*a*, *b*, *c*, and *d*).

Figure 3. Results of Simulation 2. The top panel shows the strength of the cue-outcome association in each condition. The lower panel shows the global error made by the model when attempting to infer the relative frequency of each type of trial.

Figure 4. Results of Simulation 3. The top panel shows the strength of the cue-outcome association in each condition. The lower panel shows the global error made by the model when attempting to infer the relative frequency of each type of trial.

Figure 5. Results of Simulation 4. The top panel shows the strength of the cue-outcome association in each condition. The lower panel shows the global error made by the model when attempting to infer the relative frequency of each type of trial.


0,6
0,4
0,2
0
-0,2
-0,4
-0,6
Associative strength
10
20
30
40
50
60
70
80
90
100
.20-.50
.20-.80
.50-.20
.80-.20
Trial

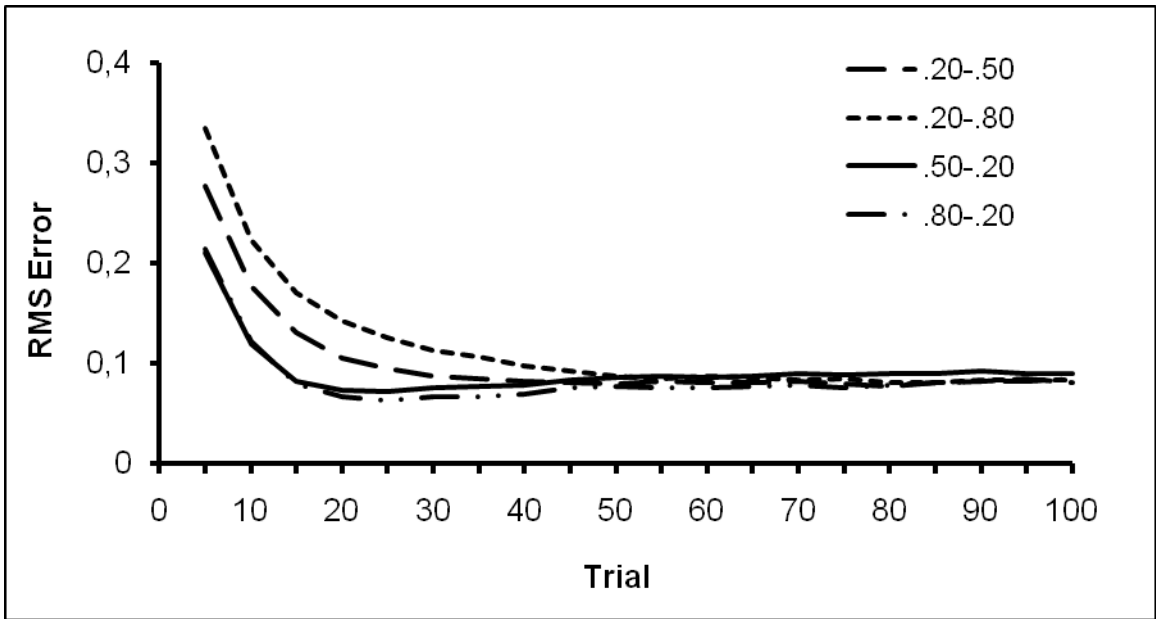

0,4
0,3
0,2
0,1
0
RMS Error
0
10
20
30
40
50
60
70
80
90
100
.20-.50
.20-.80
.50-.20
.80-.20
Trial

Figure #1

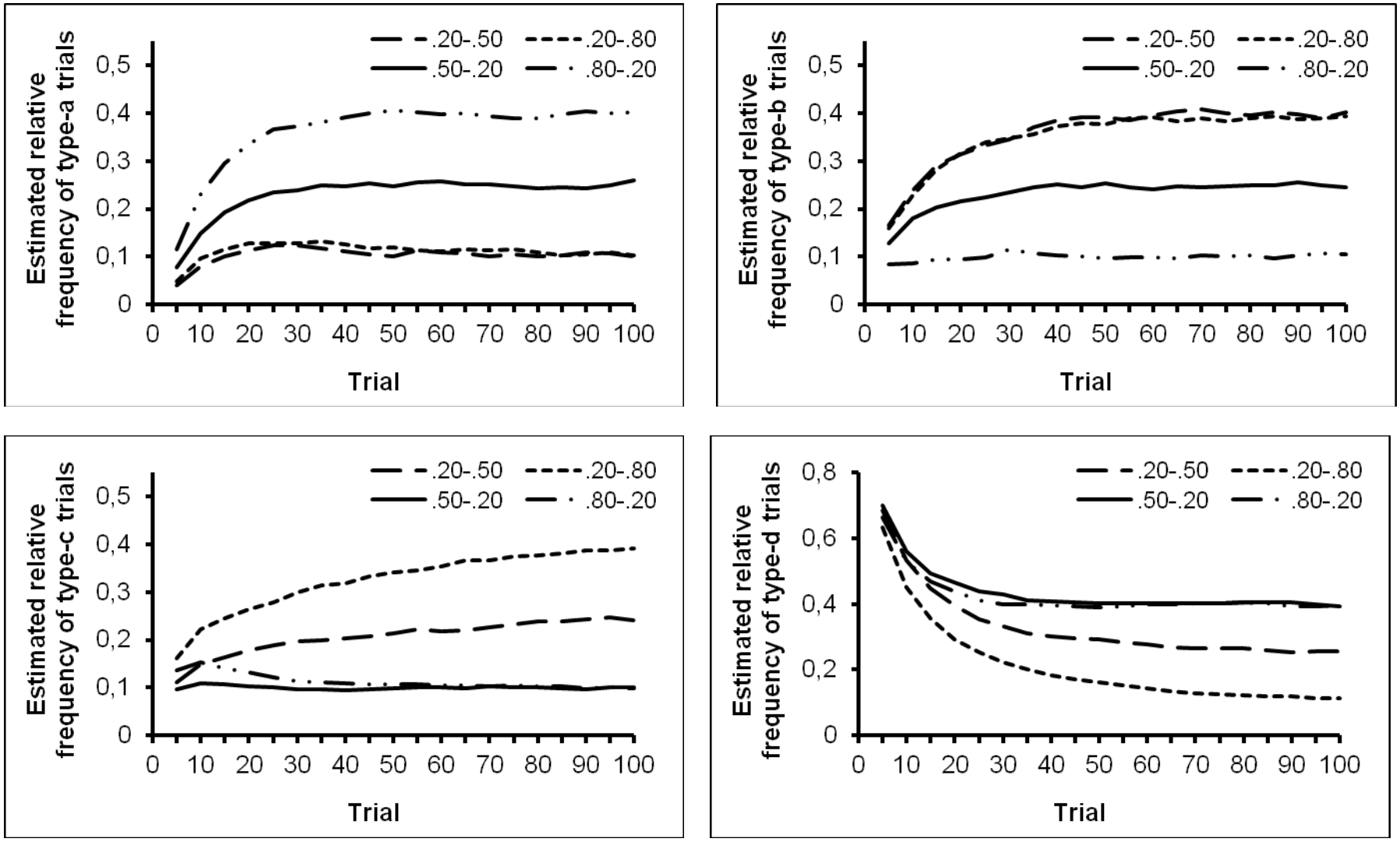
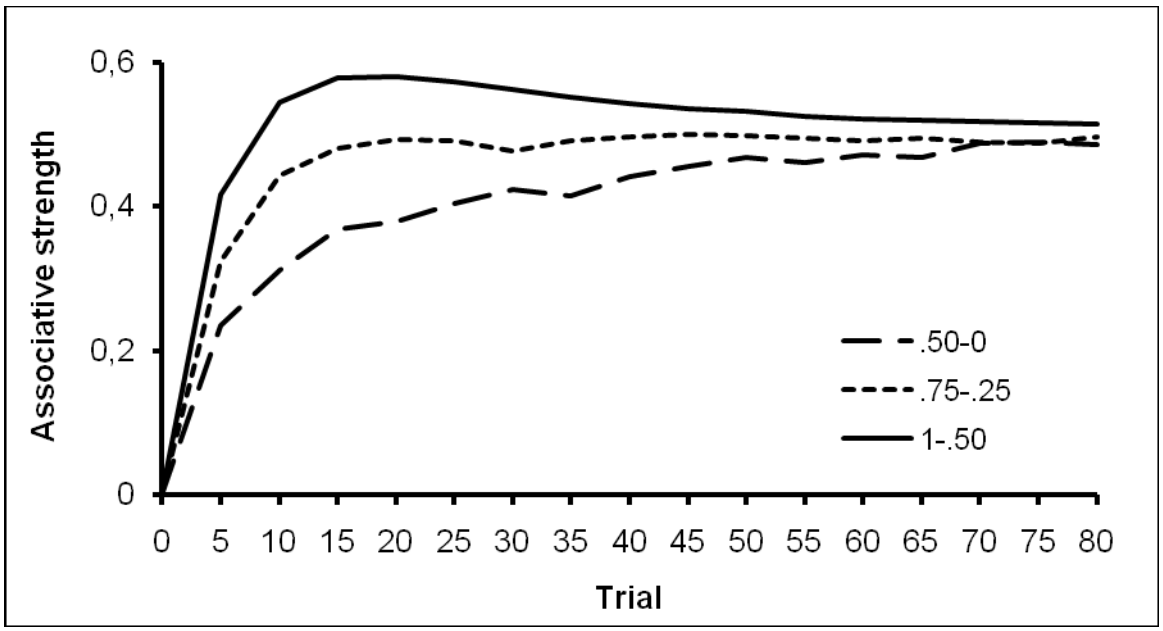

Estimated relative frequency of type-a trials
.20-.50
.20-.80
.50-.20
.80-.20
0,5
0,4
0,3
0,2
0,1
0
0 10 20 30 40 50 60 70 80 90 100
Trial
Estimated relative frequency of type-b trials
.20-.50
.20-.80
.50-.20
.80-.20
0,5
0,4
0,3
0,2
0,1
0
0 10 20 30 40 50 60 70 80 90 100
Trial
Estimated relative frequency of type-c trials
.20-.50
.20-.80
.50-.20
.80-.20
0,5
0,4
0,3
0,2
0,1
0
0 10 20 30 40 50 60 70 80 90 100
Trial
Estimated relative frequency of type-d trials
.20-.50
.20-.80
.50-.20
.80-.20
0,8
0,6
0,4
0,2
0
0 10 20 30 40 50 60 70 80 90 100
Trial

Figure #2


Associative strength
0,6
0,4
0,2
0
0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80
Trial
.50-0
.75-.25
1-.50

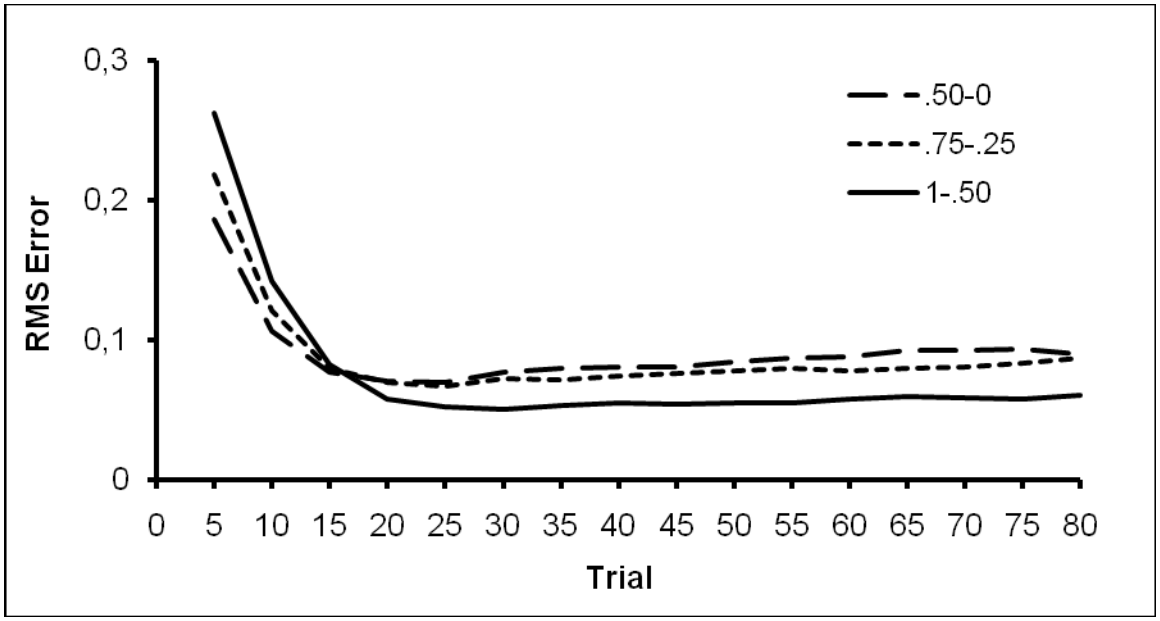
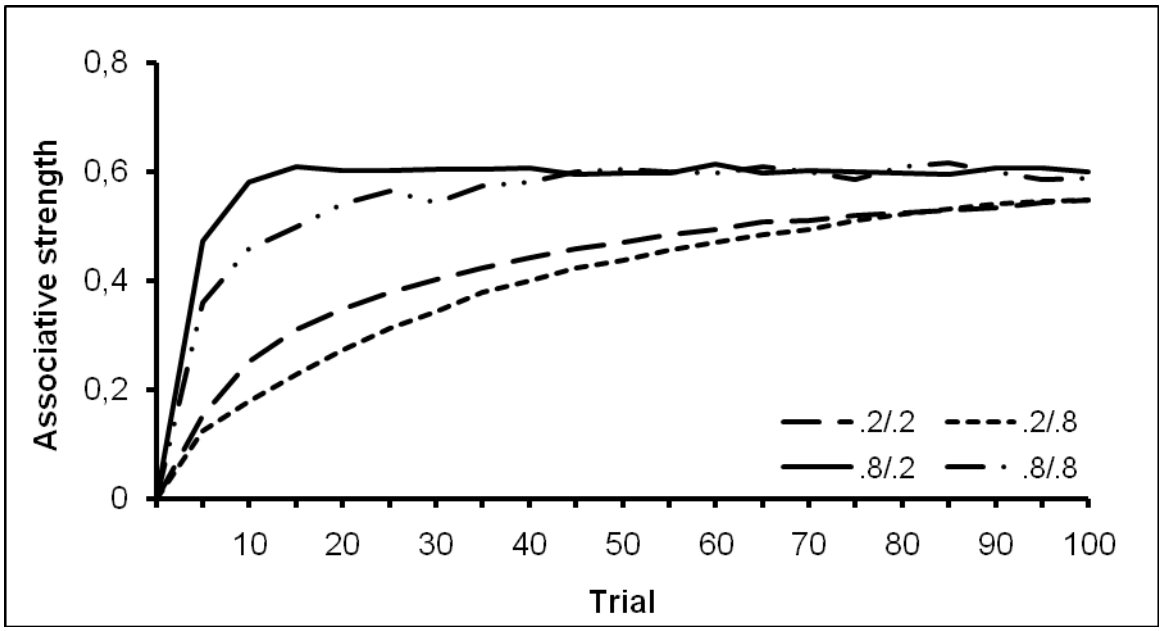

RMS Error
0,3
0,2
0,1
0
0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80
Trial
.50-0
.75-.25
1-.50

Figure #3


0,8
0,6
0,4
0,2
0
Associative strength
.2/.2
.2/.8
.8/.2
.8/.8
10
20
30
40
50
60
70
80
90
100
Trial

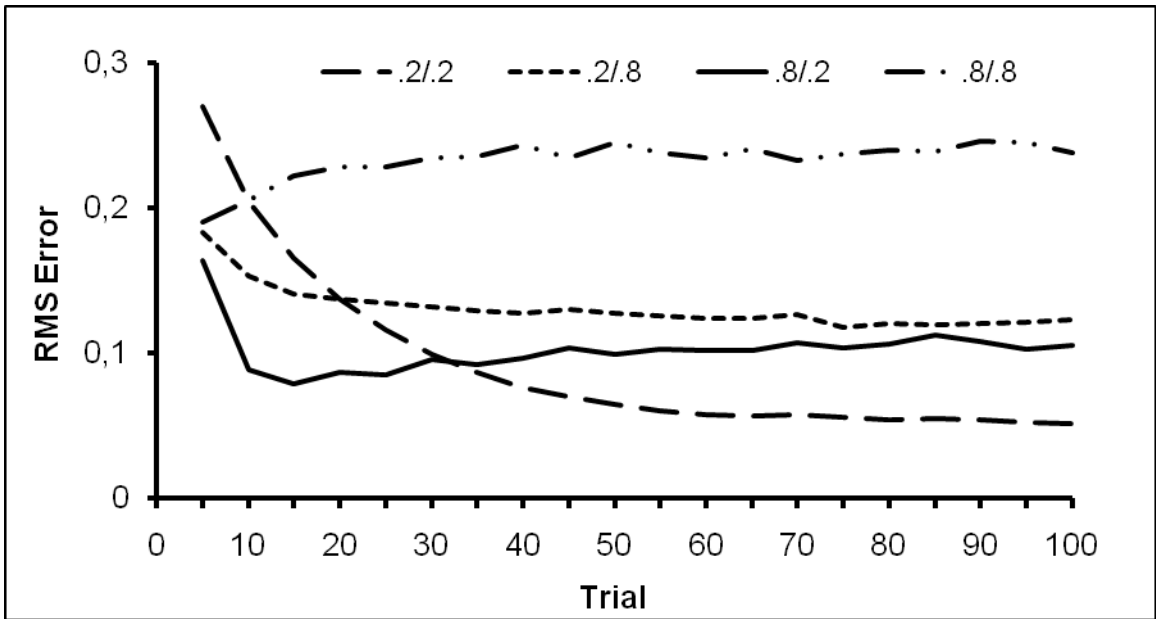
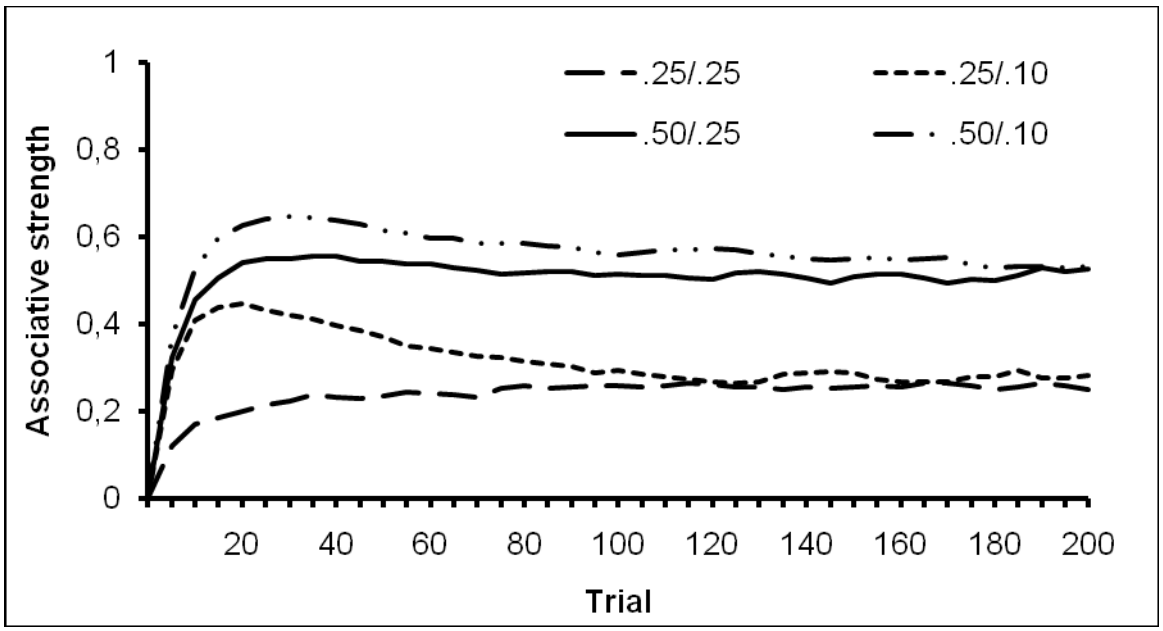

0,3
0,2
0,1
0
RMS Error
.2/.2
.2/.8
.8/.2
.8/.8
0
10
20
30
40
50
60
70
80
90
100
Trial

Figure #4


Associative strength
1
0,8
0,6
0,4
0,2
0
20
40
60
80
100
120
140
160
180
200
Trial
.25/.25
.25/.10
.50/.25
.50/.10

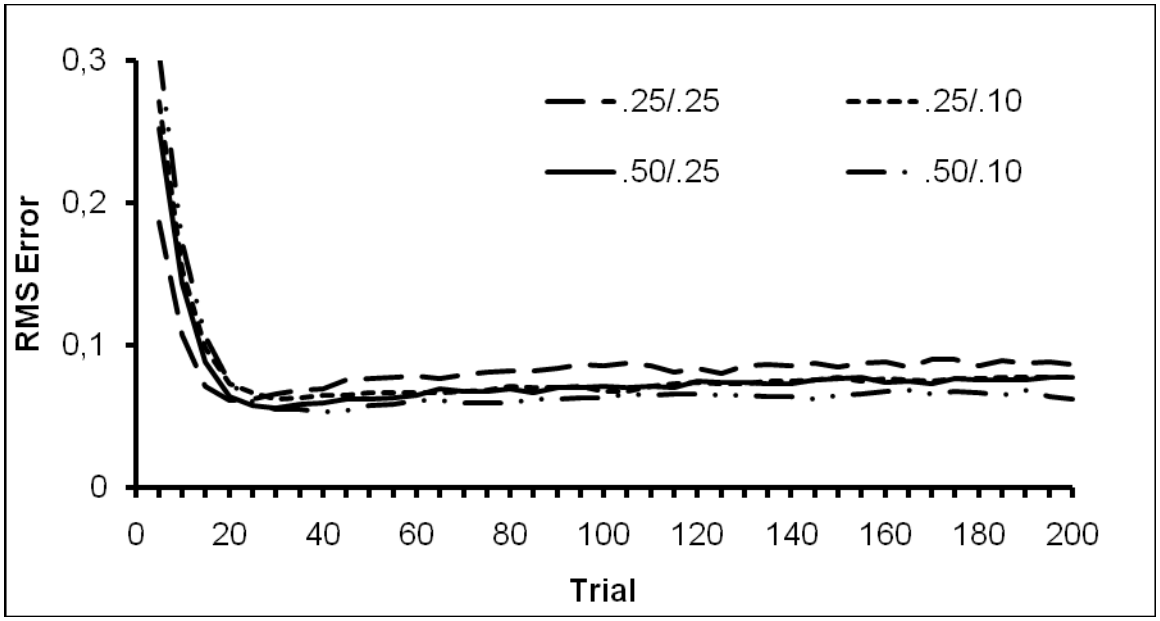

RMS Error
0,3
0,2
0,1
0
0
20
40
60
80
100
120
140
160
180
200
Trial
.25/.25
.25/.10
.50/.25
.50/.10

Figure #5