Journal of Experimental Psychology:
Animal Behavior Processes
2010, Vol. 36, No. 2, 296–306

0097-7403/10/$12.00 DOI: 10.1037/a0016603

Reasoning Rats or Associative Animals? A Common-Element Analysis of the Effects of Additive and Subadditive Pretraining on Blocking

Mark Haselgrove
The University of Nottingham

Beckers, Miller, De Houwer, and Urushihara (2006) described the results of 3 blocking experiments conducted with rats. Beckers et al. concluded that the results of these experiments cannot be accounted for with existing theories of associative learning, and argued, instead, that the results were a consequence of the rats engaging in a process akin to effortful reasoning. Simulations of the Rescorla and Wagner (1972) theory of learning presented here challenge this conclusion by providing an alternative, associative, explanation for the results presented by Beckers et al.

Keywords: blocking, reasoning, associative learning, common elements

Beckers, Miller, De Houwer, and Urushihara (2006) described the results of three animal conditioning experiments in which rats in an experimental group received paired presentations of a cue, A, with a shock before receiving paired presentations of A, in compound with X, and shock. In a final test, conditioned responding to X was weaker in this experimental group than in a control group who received an identical training history, with the exception that the initial conditioning was conducted with B rather than A.

The purpose of the experiments reported by Beckers et al. (2006) was to examine whether this so-called blocking effect, shown here using a conditioned suppression procedure in rats, was sensitive to the types of manipulations that have been shown to affect blocking of causal learning in humans. Beckers, De Houwer, Pineño, and Miller (2005, Experiment 2), for example, showed that blocking of causal learning in humans was abolished if participants first learned that stimuli C and D, and their compound CD signaled the same outcome. In addition, Beckers et al. (2005, Experiment 1) demonstrated, again using a causal learning procedure in humans, that blocking could be obtained more readily if the outcomes used in the blocking stage of the experiments were not of a maximal intensity (see also, De Houwer, Beckers, & Glautier, 2002). Beckers et al. (2005) argued that these results support the idea that blocking of human causal learning is the consequence of controlled and effortful inferential reasoning. Briefly, this reasoning is based on participants making the assumption that when two putative causes are presented in compound, as will be the case when A and X are presented in compound in a blocking experiment, a stronger outcome should occur than if only one putative cause is present (the assumption of additivity). In a blocking experiment, however, the intensity of the outcome is the same following the AX compound as when A is presented alone. Armed with this knowledge, Beckers et al. (2005) argued that people can logically deduce that X is not an effective cause of the outcome, and thus showed a blocking effect. However, if participants are given prior training that contradicts the assumption of additivity then blocking should not be observed (Beckers et al., 2005, Experiment 2). Furthermore, blocking should be facilitated by providing participants with the opportunity to appreciate that the intensity of the outcome following A is not maximal, thus demonstrating to them that X is ineffective as a cause of the outcome during the subsequent AX trials (Beckers et al., 2005, Experiment 1). In brief, the experiments reported by Beckers et al. (2006) did replicate, using a shock conditioning procedure in rats, the types of effects seen in human participants, leading them to conclude that the humble laboratory rat must be endowed with "remarkable cognitive abilities" (p. 100) and furthermore that existing theories of associative learning, many of which have been developed to account for cue-competition effects like blocking, are unable to account for these results. In fact, Beckers et al. claimed that

> Existing associative theories are actually silent when it comes to effects of additivity versus subadditivity pretraining because they do not allow abstract learning about the rules that apply to one set of cues to transfer to an entirely different set of cues. (p. 99)

Unfortunately, Beckers et al. (2006) did not demonstrate that the cues used in pretraining and those used for blocking were represented by the rats as entirely different stimuli. In fact, of the six cues that were used in these experiments, five of them were drawn from the same, auditory modality. It is conceivable therefore that the cues used for pretraining and the cues used for blocking were, to a greater or lesser degree, similar. For example, the cues used by Beckers et al. may have had a representational element in common. This being the case, then existing associative theories are not silent about the effects of additivity and subadditivity pretraining on blocking, for now the conditioned properties that are acquired to the pretraining cues can transfer to the blocking cues via this

Mark Haselgrove, School of Psychology, The University of Nottingham.

This work was supported by a grant from the Biotechnology and Biological Sciences Research Council of the United Kingdom. I am grateful to Mike LePelley for his comments on a draft of this paper and to Evan Livesey for his very helpful suggestions.

Correspondence concerning this article should be addressed to Mark Haselgrove, School of Psychology, The University of Nottingham, University Park, Nottingham, NG7 2RD England. E-mail: Mark.Haselgrove@Nottingham.ac.uk

common element. The purpose of the remainder of this article is to describe a way in which an existing theory of associative learning, the Rescorla–Wagner (1972) model, can explain the results reported by Beckers et al. by making the simple assumption that the cues employed shared a common element.

Experiment 1—Subadditive Pretraining Prevents Blocking

The top panel of Table 1 outlines the design of Experiment 1 described by Beckers et al. (2006) in which rats in the experimental and control groups, described above, were given different types of training in Phase 1 prior to the blocking procedure. For the subadditive condition, Phase 1 consisted of trials in which cues C, D, and a CD compound were each paired 4 times with shock before the elemental and compound conditioning training in Phases 2 and 3, in which A (or B) and AX were paired with shock, 12 and 4 times, respectively. The two remaining conditions served as controls: animals in the irrelevant element condition received Phase 1 training in which C, D, and E were each paired 4 times with shock, and animals in the irrelevant compound condition received pretraining in which C and a DE compound were paired with shock 8 and 4 times, respectively. Animals in these conditions, like those in the subadditive condition, then proceeded to the elemental and compound conditioning phases of Phases 2 and 3. The left panel of Figure 1, which is redrawn from Beckers et al., shows the results of the final test phase in which X was presented in isolation. The bars represent suppression ratios, which is a measure of the extent to which the presentation of X suppressed ongoing instrumental lever pressing. With this index, lower numbers reflect stronger conditioned responding. Conditioned responding was stronger in the control groups than the experimental groups for the irrelevant element and irrelevant compound conditions—the standard blocking effect. However, there was no difference between these groups in the subadditive condition; for these groups conditioned suppression was at a similar, high level.

Table 1
Experimental Designs Used by Beckers et al. (2006)

Condition and group	Phase 1: Pretraining	Phase 2: Element	Phase 3: Compound
Experiment 1			
Subadditive			
Experimental	C+, D+, CD+	A+	AX+
Control	C+, D+, CD+	B+	AX+
Irrelevant element			
Experimental	C+, D+, E+	A+	AX+
Control	C+, D+, E+	B+	AX+
Irrelevant compound			
Experimental	C+, C+, DE+	A+	AX+
Control	C+, C+, DE+	B+	AX+
Experiment 2			
Additive			
Experimental	C+, D+, CD++	A+	AX+
Control	C+, D+, CD++	B+	AX+
Irrelevant element			
Experimental	C+, D+, E++	A+	AX+
Control	C+, D+, E++	B+	AX+
Irrelevant compound			
Experimental	C+, C+, DE++	A+	AX+
Control	C+, C+, DE++	B+	AX+
Experiment 3			
Maximal			
Experimental	++, +	A++	AX++
Control	++, +	B++	AX++
Submaximal high			
Experimental	++++, ++	A++	AX++
Control	++++, ++	B++	AX++
Submaximal low			
Experimental	++, +	A+	AX+
Control	++, +	B+	AX+

Note. A, B, C, D, E and X are conditioned stimuli; +, ++, and ++++ are weak, strong and very strong shocks respectively. See text for further details. From "Reasoning rats: Forward blocking in Pavlovian animal conditioning is sensitive to constraints of causal inference" by T. Beckers, R. R. Miller, J. De Houwer, & K. Urushihara, 2006, *Journal of Experimental Psychology, General, 135,* p. 92–102. Copyright 2006 by the American Psychological Association. Adapted with permission.

This experiment was simulated with the associative theory of learning proposed by Rescorla and Wagner (1972). According to this theory, the change in the strength of the association between an individual cue and an outcome (ΔV) is determined by the difference between the asymptote of conditioning (λ) and the sum of the associative strengths of all the cues present on that trial (ΣV), multiplied by learning rate parameters for the cue (α) and the outcome (β)—see Equation 1. The associative strength of a compound consisting of two or more cues is simply the arithmetic sum of the associative strengths of all the individual cues.

$$\Delta V = \alpha \times \beta \times (\lambda - \Sigma V). \quad (1)$$

For the experiment under consideration, it was assumed that all of the cues shared a common element, thus cues C, D, E, A, B, and X were represented as compounds *cp, dp, ep, ap, bp,* and *xp*, respectively, where *p* is an element that the experimental cues shared in common and *c, d, e, a, b,* and *x* are elements that distinguish these cues.[1] A computer simulation of the Rescorla–Wagner model was used to calculate the associative strengths of these elements throughout the experiment, and to ultimately derive the terminal associative strengths of elements *x* and *p* in the experimental and control groups for the three conditions of Experiment 1 reported by Beckers et al. (2006). For this simulation, the learning rate parameters (α) associated with all the stimuli, including the common element *p*, were set to 0.2, the learning rate parameter associated with reinforcement (β) was equal to the electric current (in mA) of the shock unconditioned stimulus (US) used in Experiment 1 reported by Beckers et al., thus β was set at 0.7. All starting associative strengths (V) of the elements were zero, and the asymptote of conditioning supportable by the reinforcer (λ) was equal to β. Consistent with the associative analysis offered for other conditioning phenomena (e.g., McLaren & Mackintosh, 2002; Pearce, George, & Aydin, 2002), on compound conditioning trials (e.g., AX) the common element was represented nonadditively (e.g., *axp*). The common element *p* entered into the

[1] As *p* is assumed to be present in all cues, it is not a configural element. (Rescorla, 1972, 1973). Configural elements are activated by a specific conjunction of cues. In addition, *p* is not a contextual cue, and therefore does not undergo nonreinforcement, during simulations of the intertrial interval.

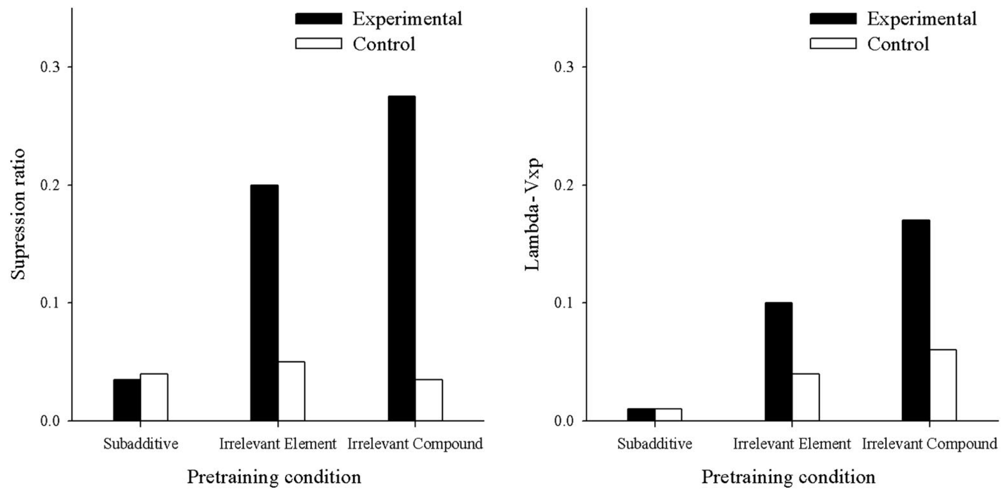

Figure 1. Left panel: The mean suppression ratios to X for the experimental and control groups for the subadditive, irrelevant element, and irrelevant compound conditions described in Experiment 1 by Beckers et al. (2006). Right panel: Output of a computer simulation of the Rescorla–Wagner model for Experiment 1 by Beckers et al. See text for further details. The data in the left panel are from "Reasoning rats: Forward blocking in Pavlovian animal conditioning is sensitive to constraints of causal inference" by T. Beckers, R. R. Miller, J. De Houwer, & K. Urushihara, 2006, *Journal of Experimental Psychology, General, 135,* p. 96. Copyright 2006 by the American Psychological Association. Reprinted with permission.

equations specified by Rescorla and Wagner (1972) in exactly the same way as any of the other elements; thus, element *p* competed for associative strength in the same fashion as, for example, element *x*. Finally it was assumed that learning proceeded to asymptote in each of the experimental phases. The right panel of Figure 1 shows the terminal associative strength of *xp* for each group and condition following Phase 3 of Experiment 1 reported by Beckers et al. To ease the comparison of this simulation with the results of Experiment 1 reported by Beckers et al.—which were expressed as suppression ratios—the output of this, and subsequent, simulations has been inverted by simply subtracting the associative strength of *xp* from the asymptote of conditioning used during the blocking stages of the experiment ($\lambda - V_{xp}$).

The output of this simple simulation of the Rescorla–Wagner (1972) model generates a reasonably good fit to the results of Experiment 1 reported by Beckers et al. (2006). The associative strengths of the test stimuli were higher in the control group than in the experimental group for both the irrelevant element condition and the irrelevant compound condition—indicating the presence of blocking in these two conditions. In fact, the magnitude of blocking was predicted to be more substantial in the irrelevant compound group compared to the irrelevant element group—exactly as shown in the data reported by Beckers et al. However, blocking was not present in the subadditive condition. Here the associative strength of *xp* was essentially at asymptote for both the experimental and control group. The reason why the Rescorla–Wagner model makes this final prediction is because the pretraining given to the groups in the subadditive condition should endow the element common to all the stimuli (*p*) with asymptotic associative strength. It is fairly straightforward to see why this is the case, following the pretraining phase:

$$V_c + V_p = V_d + V_p = V_c + V_d + V_p = \lambda$$

$$\therefore V_c + V_d + 2V_p = 2\lambda.$$

Furthermore,

$$(V_c + V_d + 2V_p) - (V_c + V_d + V_p) = \lambda$$

$$\therefore V_p = \lambda.$$

The training given in Phases 2 and 3 should not undermine the associative strength of element *p.* Furthermore, neither *a*, *b,* nor *x* will acquire any associative strength in Phase 3 (as they are blocked by *p*). Thus, when *xp* is presented in the test phase, a conditioned response equal to the asymptote of conditioning should be evoked in the experimental and control groups in the subadditive condition. More interesting, however, the pretraining given to the groups in the irrelevant element and irrelevant compound conditions should not endow element *p* with asymptotic associative strength. Consequently, there is scope for the acquisition of associative strength by *a*, *b,* and *x* in Phases 2 and 3 and hence, the possibility of revealing a blocking effect.

This prediction of the Rescorla–Wagner model is relatively resistant to variations in parameter values. For example, Figure 2 shows the output of two further simulations of Experiment 1 reported by Beckers et al. (2006). The details of this simulation were the same as for the previous simulation, with the exception that the learning rate parameter (α) for the common element *p* was set at 0.1 (left panel) or at 0.3 (right panel). The learning rate parameters for all remaining elements remained constant at 0.2. Although the terminal levels of associative strength of *xp* differ in these two simulations, the pattern of associative strengths is unchanged: blocking is still predicted to be present in the irrelevant element and irrelevant compound conditions, but not in the subadditive condition. Thus, even when the contribution to learning about the common element *p* is assumed to be either more or less substantial than the contribution of the noncommon elements, the Rescorla–Wagner model still successfully predicts the outcome of the results reported by Beckers et al. (2006).

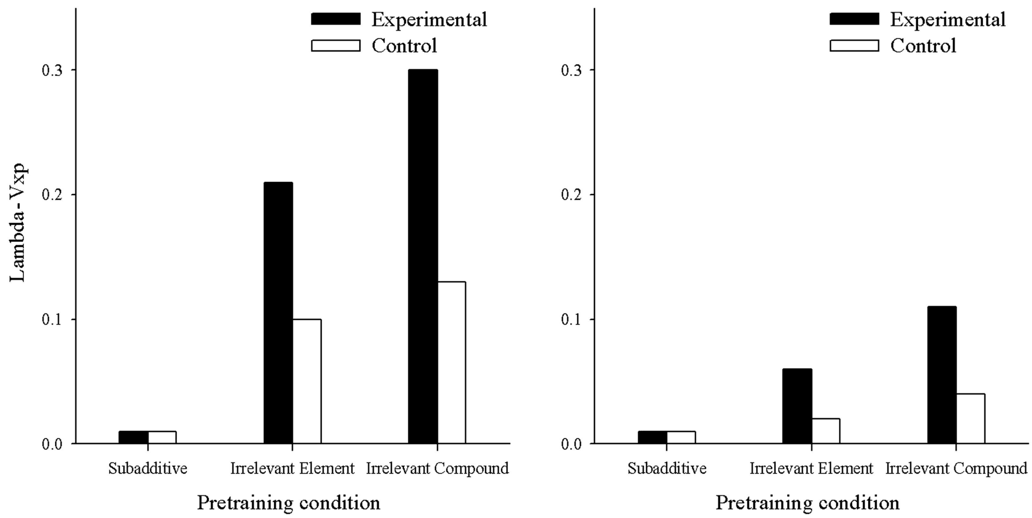

Figure 2. Output of a computer simulation of the Rescorla–Wagner model for Experiment 1 by Beckers et al. (2006). Left panel: Simulation in which the learning rate parameter for cue *p* was set at 0.1. Right panel: Simulation in which the learning rate parameter for cue *p* was set at 0.3. For both simulations, the learning rate parameters for all remaining cues was 0.2. See text for further details.

Experiment 2—Additive Pretraining Enhances Blocking

The center panel of Table 1 outlines the design of Experiment 2 described by Beckers et al. (2006) in which rats in the experimental and control groups were again given different types of training in Phase 1 prior to a blocking procedure. For the additive condition, Phase 1 consisted of trials in which C and D were each paired four times with a weak shock (0.7 mA), and a CD compound was paired four times with a stronger shock (1.0 mA) before four elemental (A+ or B+) and eight compound (AX+) conditioning trials in Phases 2 and 3. The two remaining conditions served as controls: Animals in the irrelevant element condition received Phase 1 training in which C and D were each paired with a weak shock four times and E was paired with a stronger shock, also four times. Animals in the irrelevant compound condition received pretraining in which C was paired eight times with a weak shock and a compound of D and E was paired four times with a stronger shock. Animals in these conditions, like those in the additive condition, then proceeded to the elemental and compound conditioning phases of Phases 2 and 3. The left panel of Figure 3, which is again redrawn from Beckers et al., shows the results of the final test phase in which X was presented in isolation. The bars again represent suppression ratios. Conditioned responding was stronger in the control groups than in the experimental groups for the additive and irrelevant compound conditions—indicating the presence of blocking, but there was little difference in the strength of conditioned responding between the experimental and control groups of the irrelevant element condition. The right panel of Figure 3 shows the output of a computer simulation of the Rescorla–Wagner (1972) model. The details of this simulation were identical to those used to simulate the results of Experiment 1 of Beckers et al. Thus, it was assumed that the cues used for pretraining and blocking shared a common element *p*, and that testing X was equivalent to testing *xp*. All parameters and assumptions were the same as used in the first simulation reported with the exception that both the asymptote of conditioning (λ) and the associability of the US (β) was set to 0.7 on trials with the weaker shock, and to 1.0 on trials with the stronger shock. These values are again equal to the electric current (in mA) of the shock USs used by Beckers et al. on these trials. The results of this simulation reveal that the associative strengths of the test stimuli should be substantially higher for the control groups than the experimental groups in the additive and irrelevant compound conditions—indicating the presence of a substantial blocking effect in these groups, but a notably smaller difference between the control and experimental groups in the irrelevant element condition. Like the simulation described in the previous section, this provides a reasonably good fit to the data reported by Beckers et al. in their Experiment 2.

In the simulations described thus far it has been assumed that the values of λ and β are both equal to the absolute current of the USs used in Experiment 2 reported by Beckers et al. (2006). Beckers et al., however, claimed that "psychophysically, 1.0 mA is roughly twice as aversive as 0.7 mA" (p. 97). Although Beckers et al. provided no evidence to support this claim it is possible that the use of these values for λ and β is inappropriate. Figure 4 shows the output of a series of simulations of Experiment 2 with the Rescorla and Wagner (1972) model, that were identical to the previous simulation, with the exception that the intensity and magnitude of the stronger shock was systematically varied. This was achieved by setting the value of λ and β for the weaker shock at 0.7 for all simulations and varying the value of λ and β for the stronger shock between 0.7 to 1.4. Panel A shows that when the magnitudes of the stronger and weaker shock are the same, blocking is predicted to be present in the irrelevant element and irrelevant compound groups, but not in the additive group (a result precisely that reported in Experiment 1, and the simulation for which is shown in the right panel of Figure 1). Panels B through to H show that, as

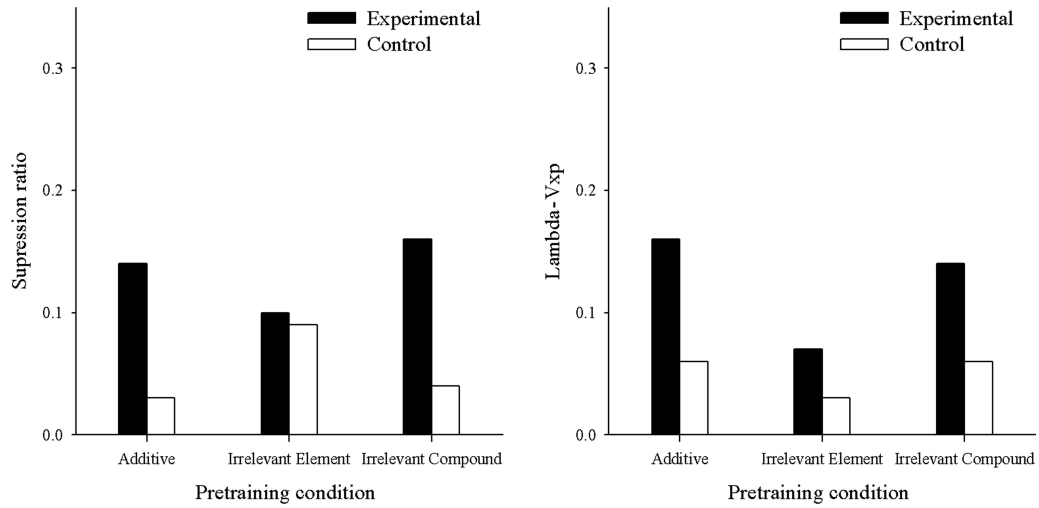

Figure 3. Left panel: The mean suppression ratios to X for the experimental and control groups for the additive, irrelevant element and irrelevant compound conditions described in Experiment 2 by Beckers et al. (2006). Right panel: Output of a computer simulation of the Rescorla–Wagner model for Experiment 2 by Beckers et al. See text for further details. The data in the left panel are from "Reasoning rats: Forward blocking in Pavlovian animal conditioning is sensitive to constraints of causal inference" by T. Beckers, R. R. Miller, J. De Houwer, & K. Urushihara, 2006, *Journal of Experimental Psychology, General, 135,* p. 97. Copyright 2006 by the American Psychological Association. Reprinted with permission.

the values of λ and β for the stronger shocks are increased above 0.7, blocking becomes progressively more apparent in the additive group, and progressively less apparent in the irrelevant element condition. There is, however, rather little evidence for variation in the degree of blocking with variation in λ and β in the irrelevant compound group. Panel H show the output of a simulation in which the values of λ and β for the stronger shock were twice that (1.4) of the weaker shock (0.7). Although, the output of this simulation still reveals the presence of blocking in the additive and irrelevant compound conditions, but not in the irrelevant element condition, the fit between this simulation and the data presented in Experiment 2 of Beckers et al. is less impressive. It must be stated though, that due to the absence of training data, it is not known whether variations (or the absence of such variation in the case of Experiment 2), in the degree of blocking between conditions, were a consequence of differences in the terminal levels of conditioned responding to the training cues (e.g., AX).

However, to reiterate a point made earlier, Beckers et al. (2006) presented no data to confirm that a shock of 1.0 mA is twice as aversive as a shock of 0.7 mA (see Campbell & Masterson, 1969, for a discussion of the psychophysics of punishment). It seems no less reasonable, therefore, to assume that the simulated values of λ and β for weaker or stronger shocks should be 0.7 and 1.0 rather than 0.7 and 1.4. In fact, given that the simulation shown in the right panel of Figure 3 used values for λ and β that were equal to the actual shock currents used in Beckers et al.'s Experiment 2, one might be more inclined to be persuaded by this simulation than others.

In any case, the account of Experiment 2 that can be derived from the Rescorla–Wagner (1972) model is more successful than the account offered by Beckers et al. (2006). Beckers et al. predicted the presence of blocking in the additive condition, but not in the irrelevant element or irrelevant compound conditions. One explanation that Beckers et al. provided for the discrepancy between the data and their predictions is that the pretraining given to animals in the irrelevant compound condition did promote the assumption of additivity. This follows because the magnitude of the outcome is greater following a compound of two cues (DE++) than during one cue alone (C+). Unfortunately, however, if this same logic is applied to Experiment 1 then the training given to animals in the irrelevant compound condition (C+, C+, DE+) should abolish blocking, as trials with a compound of two cues (DE+) and one cue alone (C+) both signal an outcome of the same magnitude. It will be recalled that blocking was observed in this condition. The simulations reported here provide a better fit to the data. As shown in Figure 4, across a range of parameters, blocking is predicted to occur following irrelevant compound pretraining, and this effect should be more substantial than the magnitude of blocking demonstrated following irrelevant element pretraining.

Experiment 3—Submaximal Pretraining Enhances Blocking

The lower panel of Table 1 outlines the design of Experiment 3 described by Beckers et al. (2006) in which rats in experimental and control groups were again given different types of training in Phase 1 prior to a blocking procedure. For animals in the maximal condition, Phase 1 consisted of pre-exposure to four unsignalled shocks of intensities 0.75mA, and 0.4 mA, before Phases 2 and 3 in which A (or B), and then AX were each paired four times with a 0.75 mA shock. Animals in the submaximal low condition received identical pre-exposure during Phase 1, before then proceeding to Phases 2 and 3 in which A (or B) and then AX were

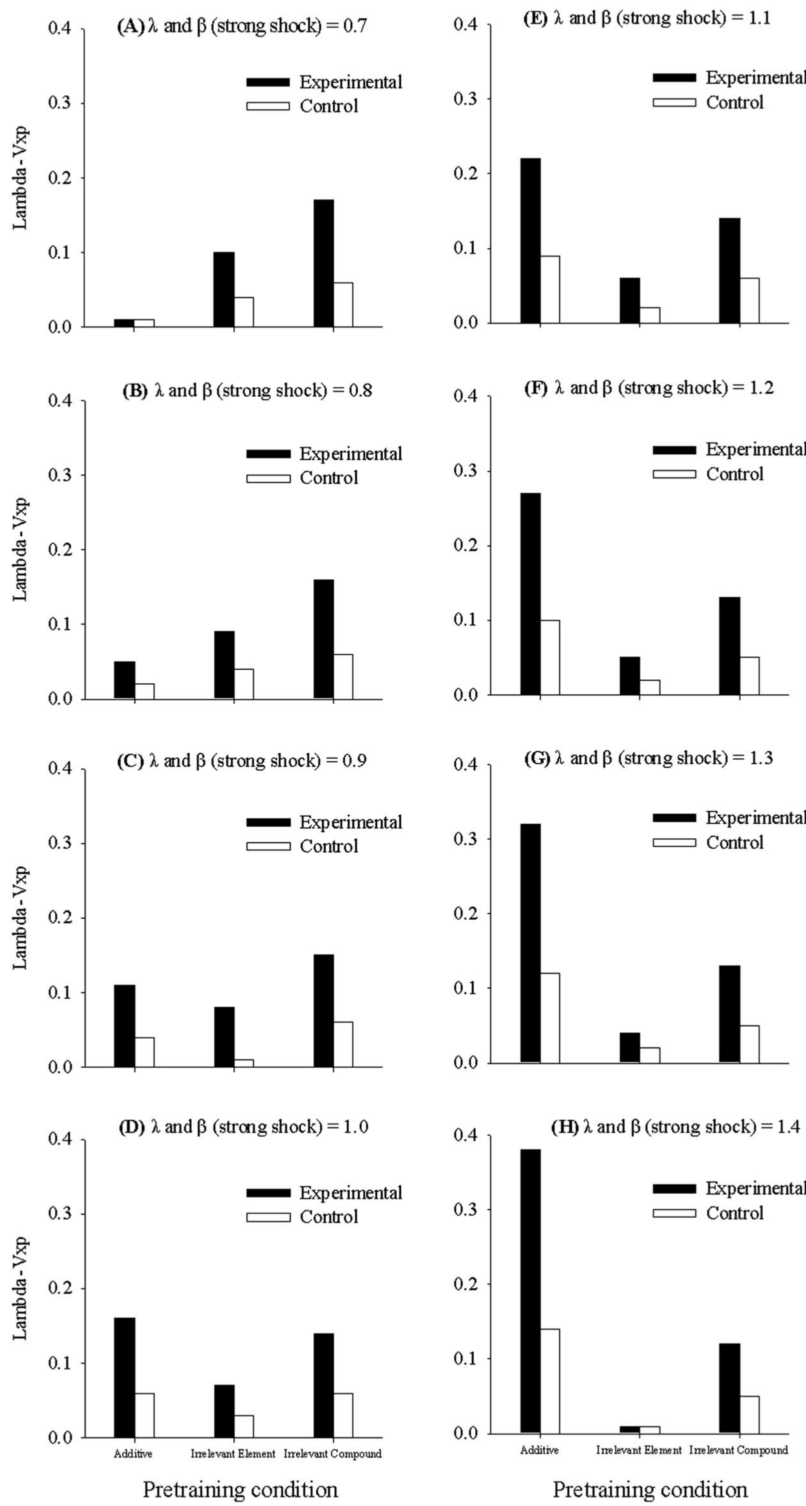

Figure 4. Output of a computer simulation of the Rescorla–Wagner model for Experiment 2 by Beckers et al. (2006). Each simulation assumed λ and β for weaker shocks = 0.7. Panels A through to H show the effects of incrementing the values of λ and β for stronger shocks from 0.7 to 1.4 in steps of 0.1. See text for further details.

paired four times each with a 0.4 mA shock. Animals in the submaximal high condition were pre-exposed to four shocks of intensities 1 mA and 0.75 mA before proceeding to Phases 2 and 3 in which they were trained in a fashion identical to the maximal condition. As in Experiments 1 and 2, all animals then received test trials with X. Figure 5, which is redrawn from Beckers et al. shows conditioned responding was weaker in the experimental groups than the control groups for the submaximal high and submaximal low conditions, indicating the presence of blocking, but there was no difference in the strength of conditioned responding in the experimental and control groups of the maximal condition. For these groups, conditioned suppression was at a similar, high level.

The critical finding of this experiment is that blocking was only obtained in the conditions in which pre-exposure was to shocks that were stronger than the shocks used for the blocking phases (submaximal low and submaximal high conditions). The presence of blocking in these conditions does not constitute a challenge to existing theories of associative learning. Even if it is assumed, as has been so far, that cues A, B, and X all share a common element, *p*, the terminal levels of conditioned responding to *xp* in the experimental groups should still be lower than in the control groups. What is more difficult to explain from just an associative perspective, however, is why pre-exposure with shocks that were no stronger than those used in the subsequent blocking training should abolish this effect. It is notable that blocking is absent in the maximal conditions by virtue of instrumental responding being almost completely suppressed. It is possible therefore that blocking persisted in this condition—as predicted by the Rescorla–Wagner model—but the effect was masked by a performance floor effect. Perhaps if the suppression of instrumental responding was given the opportunity to extinguish through, for example, a substantial

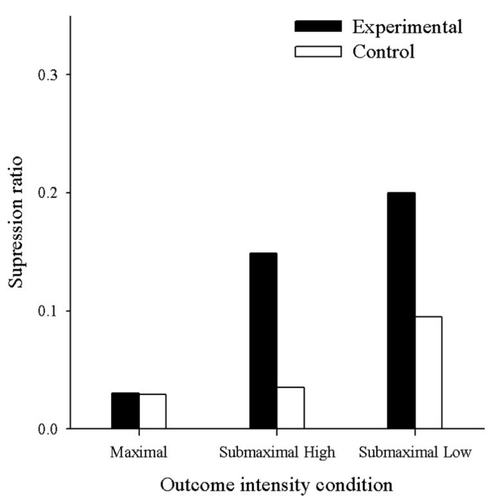

Figure 5. The mean suppression ratios to X for the experimental and control groups for the maximal, submaximal high and submaximal low conditions described in Experiment 3 by Beckers et al. (2006). From "Reasoning rats: Forward blocking in Pavlovian animal conditioning is sensitive to constraints of causal inference" by T. Beckers, R. R. Miller, J. De Houwer, & K. Urushihara, 2006, *Journal of Experimental Psychology, General, 135,* p. 99. Copyright 2006 by the American Psychological Association. Adapted with permission.

number of nonreinforced test trials with X, then blocking might be evident in the maximal group. As it was, however, X was only presented three times during the test session, which it appears was not sufficient to bring responding off the floor of performance in this condition.

This possibility clearly motivates the question of why conditioned suppression was so high in the maximal condition, which received identical element and compound training in Phases 2 and 3 as the submaximal high group and identical pre-exposure as the submaximal low group. One possibility that is worth considering is that pre-exposure in the two submaximal conditions, but not the maximal condition, resulted in a reduction in the effectiveness of the shock that was subsequently used for blocking training. This may have occurred for a number of reasons. For example, it is conceivable that the suppressive effects of the shock used during blocking was diminished by the presence of a stronger shock during preexposure.[2] Once this possibility is acknowledged it then follows that the asymptote of conditioning (λ) and perceived intensity of the US (β) may be considerably higher in the maximal condition than in either of the submaximal conditions. Consider now simulating the elemental, compound, and test stages of a blocking experiment under these circumstances, again with the assumption that the experimental stimuli share a common element, *p* (i.e., *ap*+, *axp*+, test *xp* vs. *bp*+, *axp*+, test *xp*). Although blocking will be revealed under circumstances in which λ and β are either high or low, the terminal associative strengths of *xp* following the simulation of blocking with high values of λ and β (as may be the case in the maximal condition) will be higher than the associative strengths of *xp* if rather lower values of λ and β are used (as may be the case in the submaximal conditions) and limits to performance may preclude observing differences between blocking and control groups at relatively high associative strengths. Figure 6 shows the associative strengths of *xp* following either experimental (*ap*+, *axp*+) or control training (*bp*+, *axp*+) that follow from the equations provided by the Rescorla–Wagner model. For the left pair of bars, the values of λ and β were both 0.75 (which is equal to the current [in mA] used for the shocks in the blocking stages of the maximal condition). For the center and right pair of bars the values of 0.375 and 0.2 were used respectively for the values of both λ and β. These two values represent an attempt to simulate the diminished suppressive effects of the shocks used in the submaximal high and submaximal low groups, and are simply the values of the current of the shocks (in mA) used during the blocking stages in these two groups, divided by 2. If it is assumed that limits to performance preclude the expression of variations in associative strength above a particular level (Dickinson, 1980, pp. 128–129)—in this instance, for example, $Vxp \approx 0.3$—then, as there is no overlap between the associative strength of *xp* following maximal pretraining and the associative strength of

[2] This notion is similar to suggesting that animals in the submaximal groups suffered the effects of successive negative contrast. That is, pre-exposure to a stronger shock than that used during conditioning had a negative influence on that shock's ability to subsequently serve as a US (see Mackintosh, 1974; pp. 348–402, for a discussion of contrast effects generally; Nation, Wrather, & Mellgren, 1974, for a demonstration of negative contrast effects using shocks).

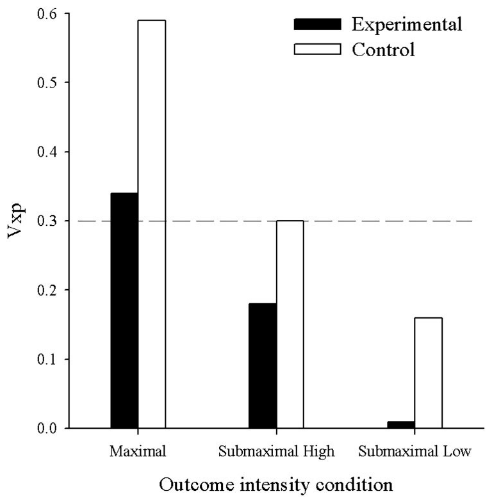

Figure 6. Output of a computer simulation of the Rescorla–Wagner model for Experiment 3 by Beckers et al. (2006). Note the ordinate now represents Vxp, not Lambda-V*xp*. The dashed line represents a hypothetical ceiling of performance. See text for further details.

xp in the remaining groups, blocking will only be observed in the submaximal conditions, but not the maximal condition.[3]

Although it is not possible to fully evaluate this explanation without further experimentation, it is notable that suppression was generally less in both submaximal conditions than in the maximal condition an effect that is consistent with the explanation offered here. Overall, however, there are no data from Experiment 3 reported by Beckers et al. (2006) that can rule out the contribution of performance effects. It therefore remains to be determined whether associative theories are challenged by experiments that demonstrate in animals an effect of maximal and submaximal pre-exposure on blocking.

General Discussion

Beckers et al. (2006) described three blocking experiments with rats, the results of which, they argued, support the idea that rats engage in a process akin to controlled, effortful reasoning and, furthermore, that existing theories of associative learning are silent with respect to these experiments. It is argued here that associative theories are far from silent in this respect. By making the assumption that cues share an element (or elements) in common, a number of the effects reported by Beckers et al. can be accommodated by the Rescorla–Wagner (1972) model. It should be noted that this assumption is by no means novel: A number of authors have considered how an element common to individual cues can influence associative learning (e.g., McLaren & Mackintosh, 2000, 2002; Pearce, George, & Aydan, 2002; Rescorla, 1976). Therefore, the simulations reported here should not be construed as a new model of learning—but instead as a common way of applying a popular model of associative learning to cues that are in some way similar to one another.

Any one of a number of different stimulus features could have served as the common element, or elements, among the stimuli in the experiments reported by Beckers et al. (2006). For example, all of the stimuli were of the same duration (30 s), it is therefore possible that the temporal properties of the stimuli served as a common feature, permitting generalization among them. All of the stimuli were presented and trained in the same conditioning chamber. It is therefore conceivable that the presentation each stimulus within the conditioning chamber activated a common representational element that was specific to that context (e.g., Wagner, 2003), or that the onset of the cues elicited a pattern of behavior specific to the conditioning chambers that mediated generalization among the cues. Whatever the nature of the common element or elements that is proposed here to be the basis of the effects reported by Beckers et al., it is reassuring to learn that when steps are taken to make the pretraining cues or context less similar to the blocking cues or context, then the attenuation of blocking seen following subadditive pretraining effect is lost (Wheeler, Beckers, & Miller, 2008). One might assume that a consequence of reducing the similarity of the pretraining and blocking cues (or the context in which these cues are presented) is to somehow compromise the associative properties that these elements share in common. A way of instantiating this is to assume the existence of not one, but several elements (e.g., *p, q, r, s*) that the pretraining and blocking cues share in common. If only a proportion of these common elements that gained associative strength during subadditive pretraining were also activated during blocking training, then the associative strengths of the blocking cues, plus the activated common elements will not be at asymptote from the outset of Phase 2 training. Consequently there is scope for variations in the associative strength of A, B, and X and a blocking effect can emerge.

So which explanation of blocking in rats should we accept—the rational or the associative? One way of addressing this question would be to exploit the fact that the associative interpretation of Experiment 1 by Beckers et al. (2006) developed here, predicts that the associative strengths of the stimuli in Phase 2 for the subadditive condition should be substantially higher than the associative strengths of the stimuli in the same phase for the irrelevant element and irrelevant compound conditions. Thus, as far as the associative account developed here is concerned, conditioned responding should be higher to A (and B) following subadditive training than following irrelevant element or irrelevant compound pretraining (but see McLaren & Mackintosh, 2002, pp. 178–182, for a discussion of how generalization of associative strength may be substantial, but may not necessarily translate into responding). Unfortunately, as was noted earlier, all of the training from Phase 2 of this experiment was conducted in the absence of any levers and therefore any data. It is not possible, therefore, to asses this prediction without further experimentation. An alternative possibility would be to explore whether additive, irrelevant element, and irrelevant compound pretraining are differentially sensitive to the difference between the magnitude of the weak and stronger USs (see Figure 4). Both the additive and irrelevant element conditions show a substantial sensitivity to varying the difference in the magnitude of the USs used during pretraining. This is not the

[3] Recent work by Rescorla (2002) highlighted the fact that it is hazardous to assume a perfect linear relationship between conditioned responding and associative strength. The performance threshold suggested here might alternatively be expressed as suggesting that differences in associative strength are more difficult to detect at higher, or lower, levels of performance.

case for the irrelevant compound condition however, which persists in showing blocking across a variety of parameters. Finally, it might be interesting to examine the logical opposite of additive pretraining—namely subtractive pretraining in which, when presented separately, cues C and D are followed by a strong US, but when presented as a CD compound, are followed by a weaker US. A simulation with the parameters and procedures used in the preceding simulations reveals that subtractive pretraining should result in superconditioning, which should be particularly evident in the blocking, rather than control group. In contrast, a deductive reasoning account of the effects of subtractive pretraining predicts—if anything—that blocking should be preserved (see Mitchell, Lovibond, & Condoleon, 2005). Using a backward blocking procedure, the latter prediction has gained a measure of support (Mitchell et al., 2005). It remains to be determined whether the same holds for forward blocking.

The simulations of the Rescorla–Wagner (1972) model that were presented here assumed that the cues used in the current experiments shared an element (p) in common with each other. This assumption was motivated by the observation that of the six cues used in the current experiments, five were drawn from the same modality. Of course not all animal conditioning experiments expose rats to quite so many stimuli or indeed draw them from the same modality. Wagner and his colleagues (e.g., Brandon, Vogel, & Wagner, 2000; Myers, Vogel, Shin, & Wagner, 2001), for example, described a number of simple compound conditioning experiments with rabbits that explicitly use cues that are drawn from different sensory modalities. Under these circumstances one might not wish to assume that these cues share a common element; or if they do, one might wish to suppose that its salience is rather negligible. In any case, making the assumption that experimental cues share an element in common does not prevent the Rescorla–Wagner model from continuing to predict many other conditioning phenomena. In the discussion to Experiment 3, it was shown how the standard blocking effect can still be simulated by the Rescorla–Wagner model, despite assuming the presence of an element common to A and X. Using the same parameters and assumptions that were used in the simulations described earlier, a simulation of training A+ and B+ and then testing a compound of AB is conceived as $ap+$, $bp+$, and a test of abp. Such training, according to the Rescorla–Wagner model, results in the associative strength of a, b, and p summing to a value greater than λ. Thus, the presence of behavioral summation (e.g., Rescorla, 1997) is not compromised by assuming that cues share a common element. Further simulations have confirmed that overshadowing, conditioned inhibition (Pavlov, 1927) and relative validity (Wagner, Logan, Haberlandt, & Price, 1968) were also preserved. It is therefore not the case that the Rescorla–Wagner model as it has been described here, is just limited to explaining the results of the experiments reported by Beckers et al.: It still provides an explanation for other conditioning phenomena.

The simulations that have been presented here have also all been based on the assumption that training, in each stage, reached asymptote. It is, of course, very difficult to know whether this was actually the case in the experiments reported by Beckers et al. (2006). It is worth considering, therefore, the implications of assuming that learning, during pretraining in particular, was not complete. The left and right panels of Figure 7 show, respectively the results of simulations of Experiment 1 and Experiment 2 reported by Beckers et al. These simulations employed the same parameter values that were used in the simulations shown in Figures 1 and 3. The ordinate on these panels is a measure of the magnitude of blocking, and is simply the terminal value of Vxp in the control group minus the terminal value of Vxp in the blocking group. Higher numbers on this scale therefore represent more blocking, zero represents no blocking. The abscissa on these panels is the number of iterations of the pretraining trial types. Shown on separate lines are the results for each pretraining condition. Looking first at the left hand panel, which show simulations from Experiment 1 reported by Beckers et al., the effects of varying the amount of pretraining on the irrelevant element and irrelevant compound conditions are relatively negligible: after 2

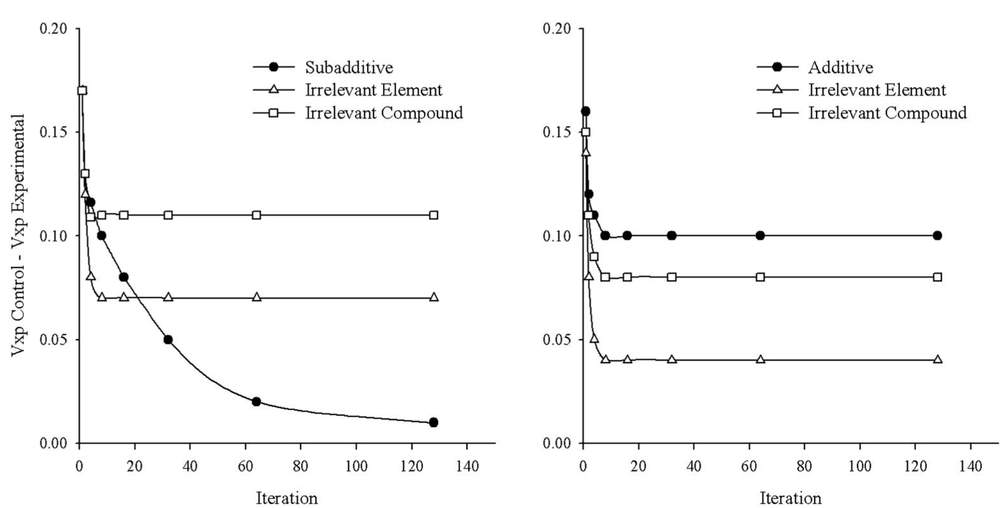

Figure 7. Output of computer simulations of the Rescorla–Wagner model for Experiments 1 (left panel) and 2 (right panel) by Beckers et al. (2006). The magnitude of blocking is plotted on the ordinate, number of iterations of pretraining is shown on the abscissa. Lines represent different pretraining conditions. See text for further details.

iterations, blocking is predicted to be more substantial following irrelevant compound than irrelevant element pretraining. There is, however, an effect on blocking of varying the amount of subadditive pretraining. Up to 8 iterations of pretraining, the magnitude of blocking following subadditive pretraining is predicted to be roughly equivalent to that following irrelevant compound pretraining. After 16 iterations, the magnitude of blocking following subadditive pretraining is predicted to be roughly equivalent to that following irrelevant element pretraining, and after this point—as asymptote is approached, the magnitude of the blocking effect diminishes. This is an interesting prediction for it does not necessarily follow from an inferential reasoning account of blocking, that limited subadditive pretraining should generate more substantial blocking than irrelevant element pretraining. It might be noted that after 4 iterations of pretraining (which corresponds to the number of pretraining trials given in Experiment 1 reported by Beckers et al., 2006) the blocking effect should have been equivalent following subadditive and irrelevant compound pretraining. However, there is no way of knowing whether four conditioning trials were equal to 4 iterations of a simulation: a single training trial may correspond to a number of iterations (e.g., McLaren & Mackintosh, 2000). There is equally no way of knowing, therefore, whether the pretraining in the experiments reported by Beckers et al. had or had not reached the point in which the present analysis predicts less blocking in the subadditive group than the others.

The simulations of Beckers et al.'s (2006) Experiment 2 that are shown in the right panel of Figure 7, however, are more robust against variations in the amount of simulated pretraining. From a simulation involving one iteration, to those at asymptote, the magnitude of blocking is consistently less following irrelevant element pretraining than following additive or irrelevant compound pretraining.

Beckers et al. (2006) provided a useful *modus tollens* argument to formally represent how an inferential reasoning account provides an explanation of blocking:

> [*if p then q*] If potential causes A and X are both effective causes of a particular outcome, then the outcome should be stronger when both are present than when only one is present.
>
> [*not q*] the outcome is not stronger when A and X are both present than when only A is present.
>
> [*therefore, not p*] Thus, A and X are not both effective causes of the outcome.

The initial training of a blocking procedure, in which A signals an outcome, establishes that A is effective in causing the outcome when presented alone. It follows, therefore, that X is not an effective cause of the outcome (Beckers et al., 2006, p. 93). If, however, the outcome were stronger when A and X are both present compared to when only A is present, then X may be concluded to be an effective cause of the outcome, and blocking should be disrupted.[4] A number of retrospective revaluation experiments contradict this prediction (Dopson, Pearce, & Haselgrove, 2009; Holland, 1999; Rauhut, McPhee, DiPietro, & Ayres, 2000; but see Blaisdell, Gunther, & Miller, 1999). For example, in an appetitive Pavlovian conditioning experiment reported by Dopson et al., rats were given A+ and then AX+ training while a control group were given A– and then AX+ training. A first test revealed stronger conditioned responding to X in the control than the blocking group. Following this successful demonstration of blocking, the blocking group were given A– training while the control group were given A+ training in a revaluation stage. A second test revealed conditioned responding to X was still stronger in the control than the blocking group. This result seems to be inconsistent with the rational inference account of blocking, for after the revaluation stage the outcome was stronger following A and X than when only A was present. Therefore, blocking should not have been apparent—yet it was.

Beckers et al. (2006) conceded that they could not "exclude the possibility that somehow an associative model could be devised that is able to account for the present results" (p. 101). The simulations described here, however, applied an existing model of associative learning (Rescorla & Wagner, 1972) to the experiments described by Beckers et al. along with the assumption that different stimuli share an element in common. This simple assumption has been widely acknowledged to provide a good account of phenomena associated with discrimination learning and stimulus generalization (e.g., Atkinson & Estes, 1963; McLaren & Mackintosh, 2000, 2002; Pearce, 1987, 1994; Rescorla, 1976) and was therefore used here to explore how a very basic theory of learning could account for the effects of additive and subadditive pretraining on blocking. The outcome of these simulations reveals that a number of the results described by Beckers et al. can be accounted for with an associative theory. Consequently, it suggests that experiments that have been thought to provide evidence for a rational account of blocking, and by implication, learning more generally, could in fact have their basis in associative learning.

[4] The structure of this argument, if *p* then *q*, *q*, therefore *p,* is a logical fallacy (affirming the consequent). However, many experiments have shown that human participants (at least) do endorse affirming the consequent as a conditional inference (Evans, Newstead, & Byrne, 1993).

References

Atkinson, R. C., & Estes, W. K. (1963). Stimulus sampling theory. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 2, pp. 121–268). New York: Wiley.

Beckers, T., De Houwer, J., Pineño, O., & Miller, R. R. (2005). Outcome additivity and outcome maximality influence cue competition in human casual learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 238–249.

Beckers, T., Miller, R. R., De Houwer, J., & Urushihara, K. (2006). Reasoning rats: Forward blocking in Pavlovian animal conditioning is sensitive to constraints of causal inference. *Journal of Experimental Psychology: General, 135,* 92–102.

Blaisdell, A. P., Gunther, L. M., & Miller, R. R. (1999). Recovery from blocking achieved by extinguishing the blocking CS. *Animal Learning and Behavior, 27,* 63–76.

Brandon, S. E., Vogel, E. H., & Wagner, A. R. (2000). A componential view of configural cues in generalization and discrimination in Pavlovian conditioning. *Behavioral Brain Research, 110,* 67–72.

Campbell, B. A., & Masterson, F. A. (1969). The psychophysics of punishment. In B. A. Campbell & R. M. Church (Eds.), *Punishment and aversive behavior* (pp. 3–42). New York: Appelton-Century-Crofts.

De Houwer, J., Beckers, T., & Glautier, S. (2002). Outcome and cue

properties modulate blocking. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 55*A, 965–985.
Dickinson, A. (1980). *Contemporary animal learning theory.* Cambridge, England: Cambridge University Press.
Dopson, J. C., Pearce, J. M., & Haselgrove, M. (2009). Failure of retrospective revaluation to influence blocking. *Journal of Experimental Psychology: Animal Behavior Processes, 35,* 473–484.
Evans, J. B. T., Newstead, S. E., & Byrne, R. (1993). *Human reasoning: The psychology of deduction*. Hove, England: Psychology.
Holland, P. C. (1999). Overshadowing and blocking as acquisition deficits: No recovery after extinction of overshadowing or blocking cues. *Quarterly Journal of Experimental Psychology, 52*B, 307–333.
Mackintosh, N. J. (1974). *The psychology of animal learning.* London: Academic.
McLaren, I. P. L., & Mackintosh, N. J. (2000). An elemental model of associative learning: I. Latent inhibition and perceptual learning. *Animal Learning and Behavior, 28,* 211–246.
McLaren, I. P. L., & Mackintosh, N. J. (2002). Associative learning and elemental representation: II. Generalization and discrimination. *Animal Learning and Behavior, 30,* 177–200.
Mitchell, C. J., Lovibond, P. F., & Condoleon, M. (2005). Evidence for deductive reasoning in blocking of casual judgements. *Learning and Motivation, 36,* 77–87.
Myers, K. M., Vogel, E. H., Shin, J., & Wagner, A. R. (2001). A comparison of the Rescorla–Wagner and Pearce models in a negative patterning and a summation problem. *Animal Learning and Behavior, 29,* 36–45.
Nation, J. R., Wrather, D. M., & Mellgren, R. L. (1974). Contrast effects in escape conditioning of rats. *Journal of Comparative and Physiological Psychology, 86,* 69–73.
Pavlov, I. P. (1927). *Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex.* London: Oxford University Press.
Pearce, J. M. (1987). A model for stimulus-generalization in Pavlovian conditioning. *Psychological Review, 94,* 61–73.
Pearce, J. M. (1994). Similarity and discrimination—A selective review and a connectionist model. *Psychological Review, 101,* 587–607.
Pearce, J. M., George, D. N., & Aydin, A. (2002). Summation: Further assessment of a configural theory. *Quarterly Journal of Experimental Psychology, 55*B, 61–73.
Rauhut, A. S., McPhee, J. E., DiPietro, N. T., & Ayres, J. B. (2000). Conditioned inhibition training of the competing cue after compound conditioning does not reduce cue competition. *Animal Learning and Behavior, 28,* 92–108.
Rescorla, R. A. (1972). "Configural" conditioning in discrete-trial bar pressing. *Journal of Comparative and Physiological Psychology, 79,* 307–317.
Rescorla, R. A. (1973). Evidence for "unique stimulus" account of configural conditioning. *Journal of Comparative and Physiological Psychology, 85,* 331–338.
Rescorla, R. A. (1976). Stimulus generalization: Some predictions from a model of Pavlovian conditioning. *Journal of Experimental Psychology: Animal Behavior Processes, 2,* 88–92.
Rescorla, R. A. (1997). Summation: Assessment of a configural theory. *Animal Learning Behavior, 25,* 200–209.
Rescorla, R. A. (2002). Effect of following an excitatory-inhibitory compound with an intermediate reinforcer. *Journal of Experimental Psychology: Animal Behavior Processes, 28,* 163–174.
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. K. Prokasy (Eds.), *Classical conditioning* (Vol. 2, pp. 64–99). New York: Appelton-Century-Crofts.
Wagner, A. R. (2003). Context-sensitive elemental theory. *Quarterly Journal of Experimental Psychology, 56*B, 7–29.
Wagner, A. R., Logan, F. A., Haberlandt, K., & Price, T. (1968). Stimulus selection in animal discrimination learning. *Journal of Experimental Psychology, 76,* 171–180.
Wheeler, D. S., Beckers, T., & Miller, R. R. (2008). The effect of subadditive pretraining on blocking: Limits on generalization. *Learning and Behavior, 36,* 341–351.

Received November 7, 2008
Revision received May 19, 2009
Accepted May 19, 2009 ■