

## Within-compound associations in retrospective revaluation and in direct learning: A challenge for comparator theory

Klaus G. Melchers and Harald Lachnit

*Philipps-Universität Marburg, Marburg, Germany*

David R. Shanks

*University College London, London, UK*

In three human causal learning experiments we investigated the role of within-compound associations in learning about absent cues versus learning about present cues. Different theoretical approaches agree that within-compound associations are essential for learning about absent cues—that is, for retrospective revaluation. They differ, however, with regard to the role of within-compound associations for learning about present cues—that is, for direct learning. A memory test was used to assess within-compound associations. Experiment 1 used a blocking/release from overshadowing design, Experiment 2 used a conditioned inhibition design, and Experiment 3 used a higher-order cue selection design. In all experiments, first-order retrospective revaluation was significantly correlated with within-compound associations, but no significant correlations were found for the direct learning conditions. In addition to this, second-order retrospective revaluation in Experiment 3 was positively correlated to joint knowledge of first-order and second-order within-compound associations. Furthermore, cue selection effects were stronger for direct learning conditions than for retrospective learning conditions. These results are at variance with the comparator hypothesis but are in agreement with a modified associative theory and with the suggestion that retrospective revaluation might be due to rehearsal processes.

*Retrospective revaluation* refers to the observation that a cue may undergo a change in its response-eliciting potential even though the cue is not present during the actual training episode. What makes retrospective revaluation so interesting is that it is at variance with traditional theories of associative learning. All of them assume that only *direct learning* can take

---

Correspondence should be addressed to Klaus G. Melchers, Psychologisches Institut, Universität Zürich, Rämistrasse 62, CH-8001 Zürich, Switzerland. Email may be sent to k.melchers@psychologie.unizh.ch or to Lachnit@mail.uni-marburg.de

The research reported in this article was supported by Grants La 564/12-1 and La 564/12-3 from the German Science Foundation (Deutsche Forschungsgemeinschaft) to Harald Lachnit and from the United Kingdom Biotechnology and Biological Sciences Research Council to David Shanks. We thank Franziska Freyberger, Anne Meinhardt, Stefanie Mooren, Sarah Schreckenberger, and Susann Wolff for their help with data collection and Ralph Miller, Helena Matute, and an anonymous reviewer for helpful suggestions on earlier versions of this paper.

The first author of this paper has previously published with “Lober” as his surname.

place, which means that changes in their response-eliciting potential should be restricted to cues that were actually present on a given trial so that their representation can be activated directly (Mackintosh, 1975; Pearce, 1994; Pearce & Hall, 1980; Rescorla & Wagner, 1972; Wagner, 1981).

Consider, for example, a standard blocking experiment (Kamin, 1968). In such an experiment, a cue A is repeatedly presented and reinforced during a first stage of learning (A+). During a second stage, A is then presented together with another cue, B, and this compound is also reinforced (AB+). When B is later tested on its own, it is usually found to evoke only a small response in comparison to a control group, which did not receive A+ training during the first stage. The most common interpretation of this finding is that A+ training results in the formation of a strong association between A and the reinforcer which then prevents or blocks the acquisition of associative strength by the redundant B stimulus during AB+ training. This is an example of cue selection where direct learning of a B–outcome association is selectively impaired by an already established A–outcome association. It is consistent with many models of associative learning. These models, however, have serious problems with observations of blocking when the two training stages are presented in reversed order (e.g., Chapman, 1991; Dickinson & Burke, 1996; Miller & Matute, 1996; Shanks, 1985). When A and B are presented and reinforced together during the first stage, they should both have intermediate levels of associative strength. A+ training during Stage 2 should then only increase A's associative strength and should not affect B retrospectively. Yet, it has been found that B's potential to elicit a conditioned response decreases as a result of a procedure in which AB+ training is followed by A+ training.

Additional empirical observations of retrospective revaluation that are problematic for associative theories include (amongst others) backward conditioned inhibition (Chapman, 1991) and release from overshadowing (Kaufman & Bolles, 1981; Matzel, Schachtman, & Miller, 1985; Wasserman & Berglan, 1998). In a backward conditioned inhibition design, simultaneous presentations of A and B without reinforcement (AB–) are followed by A+ trials that result in B gaining inhibitory properties. And in a release from overshadowing experiment, AB+ presentations are subsequently followed by an A– training stage which leads to increases in B's potential to evoke a conditioned response.<sup>1</sup>

The aim of the present research was to evaluate different accounts that have been offered to explain retrospective revaluation and to assess their predictions about the role of within-compound associations. First, we will present three of these different approaches: the comparator hypothesis (Denniston, Savastano, & Miller, 2001; Miller & Matzel, 1988), modified associative theories (Dickinson & Burke, 1996; Van Hamme & Wasserman, 1994), and the suggestion that rehearsal processes are responsible for retrospective revaluation (Chapman, 1991). Then we briefly discuss some of the empirical evidence that might allow us to

---

<sup>1</sup>It should be noted that there is also evidence of another group of revaluation effects in which there seems to be a change of the response-eliciting potential of a cue in the same direction as for a cue that is present—that is, the change would be in the opposite direction compared to the effects mentioned above. These effects are examples of so-called mediated conditioning. They are outside the scope of the theories discussed in the present paper and will receive no further attention here. Nevertheless, we want to mention that recent findings suggest that the type of reinforcer used in an experiment seems to be one of the important aspects that is responsible for the kind of revaluation effect that occurs (Dwyer, 1999).

distinguish between the comparator hypothesis and the other two accounts. Finally, we report three experiments that allow further empirical assessment.

The first account is the comparator hypothesis suggested by Miller and his colleagues (e.g., Denniston et al., 2001; Miller & Matzel, 1988).<sup>2</sup> This account does not assume that cue selection effects like blocking reflect the operation of a selective learning mechanism but, instead, assumes that they reflect effects of performance. In contrast to contemporary associative theories, the comparator hypothesis assumes that mere contiguity is sufficient for the development of an association between a cue and an outcome. Strong excitatory responding to a cue, however, will only be apparent when its association with the reinforcer is considerably stronger than the associative strength of its comparator stimulus (the learning context or another cue that was presented simultaneously with the cue in question). In the case of a blocking experiment, this means that cue B is blocked because its association with reinforcement is weaker than the association of its comparator stimulus A with reinforcement. Whether the two training stages of a blocking experiment are presented in the standard forward manner or in a backward manner should be immaterial. The only crucial factor is that the relevant comparator stimulus is activated properly. In a blocking preparation, this will be the case when cue B and its comparator stimulus A are associated with each other via within-compound associations. Those within-compound associations are built between cues that occur concurrently (Rescorla & Durlach, 1981; see Denniston et al., 2001, for a more detailed discussion of the actual comparator process and the role of within-compound associations). As a consequence, forward and backward blocking effects should both be correlated with the strength of the relevant A–B within-compound associations.

Modifications of contemporary associative theories constitute the second approach to deal with retrospective revaluation (e.g., Dickinson & Burke, 1996; Van Hamme & Wasserman, 1994). In contrast to the comparator hypothesis, they view revaluation as a learning phenomenon. Van Hamme and Wasserman (1994) introduced a modification of the Rescorla–Wagner (1972) model, and Dickinson and Burke suggested a modification of Wagner’s (1981) SOP model. These modified associative theories assume that absent as well as present cues may undergo changes of their associative strength. However, not all cues that are absent on a given trial will have their weights modified but only those that are *relevant*, which means that they are associatively activated by virtue of the fact that they have a within-compound association with the actual cue (Dickinson, 2001).<sup>3</sup> Those within-compound associations will then allow the retrieval of absent cues so that their associative strengths can be modified. The important point, now, is that the change of the associative strength of an absent cue is in the opposite direction as the change of the association of a present cue. For the case of backward blocking, this means that when A+ trials are presented during the second stage, A’s associative strength will increase while the associative strength of the absent cue B will decrease at the same time. In

---

<sup>2</sup>The comparator account of blocking and other effects of selective learning is very similar to some recent models from the field of cognitive psychology (Cheng, 1997; Cheng & Holyoak, 1995). For the present experiments, these make the same predictions as the comparator hypothesis. Therefore, all later arguments with regard to the comparator hypothesis are also valid for these cognitive models.

<sup>3</sup>Note that this description differs slightly from the first applications of the Van Hamme and Wasserman (1994) model (Van Hamme, 1994; Wasserman, Kas, Van Hamme, Katagiri, & Young, 1996). This is because the question of the relevance of a cue had not yet been solved satisfactorily in those applications.

contrast to the comparator hypothesis, within-compound associations are only necessary for retrospective revaluation but not for direct learning in the forward blocking design. Thus, backward blocking effects should be correlated with the strength of the within-compound associations while forward blocking should not be related to them.

The third account of retrospective revaluation assumes that earlier relevant trials are *rehearsed* during the second stage of a backward blocking experiment (Chapman, 1991; see also Baker & Mercier, 1989). As in the modified associative theories, relevance can be assumed to depend on within-compound associations. However, instead of modifying the associative processes that take place during learning, this approach assumes that during the second stage of a backward blocking experiment, when only A+ trials are presented, AB+ trials will be rehearsed. Thus, subjects mentally present themselves with an AB+ episode while they actually experience only an A+ trial. This has consequences similar to those of real training with AB+ and A+ trials intermixed. If one applies a competitive associative learning rule (e.g., Rescorla & Wagner, 1972) to this kind of training, then this training results in a decrease in B's response-eliciting potential while A's response-eliciting potential increases. Due to poor learning of the within-compound associations or limited working memory capacity, however, rehearsal might not always take place. As for the modified associative theories, this means that effects like blocking should crucially depend on within-compound associations in a retrospective revaluation design. In contrast to this, they do not depend on within-compound associations in conditions that only require direct learning. If A, for example, has already reached its asymptotic associative weight as a consequence of A+ training during the first stage, later rehearsal during the AB+ trials does not increase the strength of the blocking effect. Instead, rehearsal in a direct learning condition can only have an effect at all if the cues that were presented during the first stage have not yet reached their asymptotic values. In such a case, rehearsal would strengthen the blocking effect, and a weak correlation between blocking and memory for within-compound associations might result. But even then, rehearsal should not qualitatively influence whether blocking does or does not occur.

Some empirical evidence might allow a distinction to be drawn between the different accounts. Taken together, the different assumptions about the role of within-compound associations make it possible to test the comparator hypothesis against the modified associative theories and the rehearsal account. Whereas the comparator theory assumes that within-compound associations are equally relevant for both direct learning and retrospective revaluation, the later two accounts assume that within-compound associations are only crucial for retrospective revaluation; from the standpoint of these two accounts, within-compound associations should not or, at best, only be weakly correlated with cue selection effects in direct learning conditions.

A number of human causal learning studies have manipulated learning of within-compound associations and investigated the effects on direct learning and on retrospective revaluation. As it is assumed that a within-compound association between two cues is learned when they are repeatedly presented together as a compound (Rescorla & Cunningham, 1978; Rescorla & Durlach, 1981), most experiments manipulated whether or not cues were presented in consistent pairs (Dickinson & Burke, 1996; Larkin, Aitken, & Dickinson, 1998). This means that those studies manipulated whether two cues were only shown together (e.g., only AB compounds) or whether they were also presented as parts of other compounds (e.g., also AC and BD compounds). Presenting cues as parts of different compounds was assumed to

weaken the formation of within-compound associations. Furthermore, Aitken, Larkin, and Dickinson (2001) have recently used serially presented compounds for which the interstimulus interval was either unfilled or was filled with a secondary task. Again, the latter manipulation was assumed to weaken within-compound learning. The basic finding in all of these studies was that these manipulations impaired retrospective revaluation but left direct learning unaffected. These results are in agreement with modified associative theories and with the rehearsal account, but they are at variance with the comparator theory.

In a related line of research, investigators tried to assess within-compound associations by means of post-learning memory tests. Results from these studies, however, have not yielded a clear pattern. On the one hand, there is evidence that participants with better performance in the memory test show stronger revaluation effects. Chapman (1991), for example, reported a significant correlation between backward blocking and memory performance in one of her experiments. Similarly, Wasserman and Berglan (1998) and Aitken et al. (2001, Experiment 3) classified participants post-hoc on the basis of their memory performance and found that revaluation was stronger in those participants with better performance in the memory test. Furthermore, there was no evidence in Aitken et al.'s experiment that the strength of the blocking effect was related to memory performance in a condition that only required direct learning. On the other hand, the studies by Chapman (1991), Aitken et al. (2001), and by Wasserman and Berglan (1998) have some potential limitations. In Chapman's study, it is unclear how participants were "queried about their memory for which . . . [cues] had previously appeared in compound together" (Chapman, 1991, p. 843). Moreover, there is no obvious reason why she only found a significant correlation between memory performance and retrospective revaluation in one of the four experiments in which she assessed memory performance but not in the others. Evaluation of this aspect is further complicated because she did not report the exact sizes of the other correlations, which makes comparisons between direct learning and retrospective revaluation (and between different experiments) difficult. Such a comparison is also not possible on the basis of Wasserman and Berglan's (1998) data, because they only tested retrospective revaluation conditions. The data from Aitken et al.'s (2001) experiment partly allow such a comparison, as both conditions were included. Their classification, however, was based on whether participants remembered more compounds that were used for the retrospective revaluation condition or more that were used for the forward blocking condition. Evidence for retrospective revaluation was only found for those participants who remembered more revaluation compounds but not for those who remembered more blocking compounds, whereas forward blocking was not related to this classification. Nonetheless, one might expect that memory for revaluation compounds should correlate with the size of the revaluation effect in both groups.

Taken together, studies that manipulated the learning of within-compound associations indicate that these are crucial for retrospective learning but do not influence direct learning. Yet the evidence from experiments that tried to assess within-compound associations with memory tests seems less convincing with regard to their role. Therefore, the aim of the present research was to assess the degree to which retrospective revaluation and direct learning are related to within-compound associations. In contrast to the work by Dickinson and his collaborators (Aitken et al., 2001; Dickinson & Burke, 1996; Larkin et al., 1998), we did not try to impair learning about these associations but tried to assess their role when no additional manipulations were included. To assess within-compound learning, we used a memory test

like the one used in their studies. Experiment 1 used a combined blocking and release from overshadowing design, and Experiment 2 used a conditioned inhibition design. With Experiment 3, we then extended our investigation to higher-order cue selection designs.

## EXPERIMENT 1

We used a causal learning scenario that asked participants to imagine that they were allergists who wanted to find out which of a variety of different foods cause allergic reactions in one of their patients. Thus, the various foods were the causal cues, and the allergic reaction was the outcome. Participants were told that the learning trials were successive days on which the patient had eaten either one or two kinds of food and that their task was to find out which foods caused an allergic reaction in the patient.

Participants had to rate the causal relationship of the different cues with regard to the occurrence of an allergic reaction twice: the first rating had to be given before the start of the experiment and the second after the second learning stage had finished. Furthermore, we used a recognition test after the second rating to assess participants' knowledge of within-compound associations.

Table 1 illustrates the design of Experiment 1. For each participant we assessed both retrospective learning and direct learning conditions. For the retrospective revaluation conditions, during Stage 1 cues A and B as well as C and D were presented together as compounds and were always followed by an outcome (AB+, CD+). During Stage 2, A and C were presented on their own. A was always reinforced, and C was always nonreinforced. The AB+, A+ trials thereby constitute a backward blocking design, and the CD+, C- trials constitute a release from overshadowing design. Retrospective revaluation might occur for both pairs of cues, with B losing its response-eliciting potential as a consequence of the A+ trials and D gaining response-eliciting potential as a consequence of the C- trials. Cues E to H were used to set up the same cue contingencies in a direct learning condition. In Stage 1, single cue presentations were shown (E+, G-); compound cues were shown in Stage 2 (EF+, GH+). Additional filler trials ensured that participants also experienced nonreinforced compound trials during both

TABLE 1  
Design of Experiments 1 and 2

<i>Experiment</i>	<i>Condition</i>	<i>Stage 1</i>	<i>Stage 2</i>	<i>Memory test</i>
1	Retrospective revaluation	AB+, CD+	A+, C-	A?, C?
	Direct learning	E+, G-	EF+, GH+	E?, G?
	Filler	IJ-, K-	I-, KL-	I?, K?
2	Retrospective revaluation	AB-, CD-	A+, C-	A?, C?
	Direct learning	E+, G-	EF-, GH-	E?, G?
	Filler	IJ+, K+	I+, KL+	I?, K?

*Note:* A to L are different foods, "+" represents the presence of an allergic reaction, and "-" represents the absence of an allergic reaction. Each trial was presented eight times, and each participant worked on two replications of this design (i.e., two retrospective revaluation conditions, two direct learning conditions, and two sets of filler trials). "?" means that participants were asked with which other cue the given cue had previously been presented. Before Stage 1 and after Stage 2, participants had to rate all the different cues.

stages and that the numbers of reinforced and nonreinforced trials in each stage were the same. Each participant worked on two concurrent replications of the design described in Table 1. This means that two different pairs of cues were used for each condition, so that participants were trained with 24 food cues altogether (i.e., two times the 12 cues from Table 1). The trials for the two replications were presented interspersed.

## Methods

### *Participants*

The participants were 36 University of Marburg students, 12 were male and 24 were female; their ages varied between 19 and 47 years, with a median of 22.5. They either received course credit or were paid for their participation. Participants were tested individually and required approximately 30 min to complete the experiment.

### *Apparatus and procedure*

The instructions and all necessary information were presented on a PC computer screen, and the participants gave their answers by using the mouse or the computer keyboard.

The following foods were used as cues for the experiment: apples, avocados, bananas, blueberries, broccoli, carrots, cherries, coffee, eggs, fish, grapes, ice-cream, kiwi fruits, lemons, meat, mushrooms, nuts, pears, peppers, popcorn, potatoes, strawberries, toast, tomatoes. Counterbalancing ensured that each food was used equally often for each kind of cue. That means that across participants each food was used three times as an A cue either in the design in Table 1 or in the replication, three times as a B cue, three times as a C cue, and so on.

At the start of the experiment each participant was shown the following instructions (in German) on the screen:

This study is concerned with the question of how people learn about relationships between different events.

In this experiment, you will be asked to judge the probability that certain foods lead to the occurrence of allergic reactions. For this reason, you should first give your current views on different foods. This is to check whether you suspect that certain foods that are used in this experiment are particularly associated with allergies.

Each food will be presented on the screen by itself. Using a scale from 0 to 10, you should then indicate to what degree you believe that this food leads to the occurrence of an allergic reaction in a typical person. This scale will also be shown on the screen. Give your judgement by moving the slider of the scale with the mouse or by clicking on the slider if you want to retain the default value. After giving a judgement for a food, you can click on the “next” button to move on to the next food.

For your judgement, please imagine that an ordinary person who is unknown to you has eaten the specified food. On the scale, please indicate how probable you think it is that the given food leads to an allergic reaction.

As most people do not suffer from allergies, you may feel that the imaginary person will show an allergic response to none of the presented foods.

The purpose of the present rating is to see whether you suspect that certain foods particularly cause allergies, which might then influence your later judgements.

After clicking on the NEXT button, participants had to rate the causal relationship between each food and the occurrence of an allergic response in general. We used a scale that ranged from 0 (*does not cause an allergic reaction*) to 10 (*causes an allergic reaction*). Participants had to give their ratings by moving a slider on this scale. Values between 0 and 10 were not labelled, but the respective values appeared next to the slider when it was moved. The initial position of the slider was set at a value of 5. Each cue was presented on its own for this rating procedure. The 24 (8 retrospective reevaluation, 8 direct learning, 8 filler) cues were rated in random order.

Participants then received another window with the following instructions:

Now imagine that you are a medical doctor. One of your patients often suffers from allergic reactions after meals. To discover which foods lead to allergic reactions, your patient eats specified foods on each day and observes whether or not an allergic reaction occurs.

The results of these daily allergy tests are shown to you on the screen one after the other. You will always be told what your patient has eaten on a given day. On some days, he has only consumed a single kind of food, and on some days he has consumed two different foods. Please look at the foods carefully.

Then you will be asked to predict whether or not the patient will show an allergic reaction to the food he consumed. To make this prediction, please click on the appropriate response button. After you have made your prediction, you will be informed whether or not your patient actually showed an allergic reaction. Use this feedback to find out to which foods your patient reacts allergically.

Obviously, at first you will have to guess, as you do not know anything about your patient. But eventually you will learn which foods cause allergic reactions in this patient and will be able to make correct predictions.

Later in the experiment, you will be asked to judge again the probability with which the different foods cause allergic reactions. For this later judgement, however, you should indicate the probability with which your specific patient reacts allergically to a certain food. Use all of the information that you have collected during the daily allergy tests up to that time. For all of your answers, accuracy rather than speed is essential. Please do not take any notes during the experiment.

If you have any more questions, please ask them now. If you don't have any question, please start the experiment by clicking on the "next" button.

When a participant asked a question, it was answered by the experimenter. After clicking on the NEXT button, the learning stages started. Both stages consisted of eight blocks. For both replications, every trial type was presented once in each block. Thus, a total of 192 trials were presented to the participants. The order of presentation of the trials within each block was determined randomly for each block and each participant.

On each learning trial, pictures of one or two cues were shown on the screen. These were accompanied by their respective names. Single cues were presented in the centre of the program window, and compound cues were presented side by side, with left-right allocation determined randomly on each trial. Participants were told that their patient had eaten those foods. On the basis of those cues, participants had to make a prediction of whether or not they expected that an allergic reaction would occur. This was done by clicking one of two answer buttons, which were labelled "Yes, I expect an allergic reaction" and "No, I do not expect an allergic reaction". After making a prediction, another window appeared, which told participants whether or not their patient had shown an allergic reaction. Depending on the actual outcome, different colours were used for this feedback (red for the occurrence of an allergic response and green for non-occurrence). Participants had to confirm that they had read the feedback by clicking on an "OK" button. Then the next trial started.

After Stage 2, participants received the following instructions for the causal rating test:

Now please judge the probability with which the different foods lead to allergic reactions in your patient.

Use all of the potential information that you have collected about your patient.

The rating procedure was identical to the procedure at the start of the experiment with the exception that participants now had to rate the causal relationship between each food and the occurrence of an allergic response specifically for their patient.

After the second rating phase, a recognition test was used to assess knowledge of within-compound associations. Each cue that had been presented on its own during the experiment (A, C, E, G, I, and K and the cues of the respective replications) was presented alone, and participants had to say with which other cue (shown on a list with all 24 cues) it had appeared together as a compound. As for the rating phase, order of presentation was determined randomly for the recognition test. For this recognition test, the participants were shown another screen with instructions:

During the experiment your patient sometimes consumed two foods on the same day. Now it is your task to recall which foods were consumed together.

For this purpose, you will be shown one food at a time on the screen. Please remember which other food your patient ate together with that food.

To assist you, you have two sheets on which are shown all the foods that your patient ate.

These sheets are in the folder in front of you on the table. The foods are ordered alphabetically and numbered on these sheets.

For your answers, please give the number of the food that your patient ate together with the food presented.

If you feel unsure which foods were eaten together, please try to guess as accurately as possible.

### *Data scoring and aggregation*

The trial by trial predictions of the two replications of each type of trials (i.e., AB+ and its replication, CD+ and its replication etc.) were averaged for each block. Likewise, the causal ratings for the different cues for the two replications were averaged for each type of cue.

## Results and discussion

The acquisition data can be seen in Figure 1, which shows the percentage of trials for which participants predicted an allergic reaction. For both learning stages, it is obvious that these predictions closely followed the actual contingencies as early as from Block 2 onwards. This means that reinforced cues received a high percentage of allergy predictions and nonreinforced cues a low percentage. All participants met the learning criterion of more than 70% correct predictions during the second half of each stage. During these second halves, participants' predictions were correct on 98.26 and 97.80% of the trials in Stage 1 and Stage 2, respectively. Thus, participants successfully learned which cues or cue compounds predicted an allergic reaction.

We compared the ratings of the different cues that were made after the second learning stage by a one-way repeated-measures analysis of variance (ANOVA). This ANOVA revealed a significant effect of cue,  $F(11, 385) = 179.08$ ,  $MSE = 3.16$ ,  $p < .001$ . An additional ANOVA

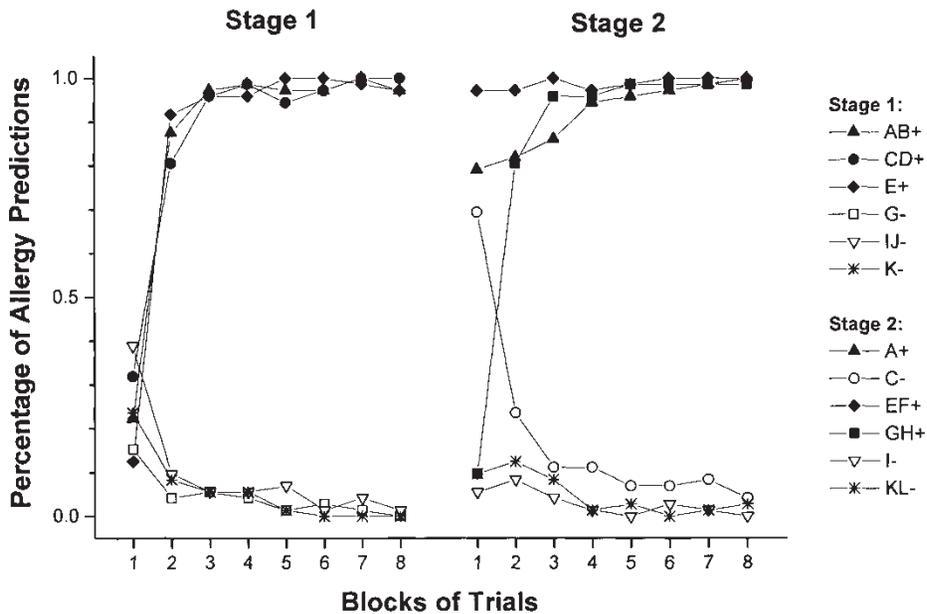
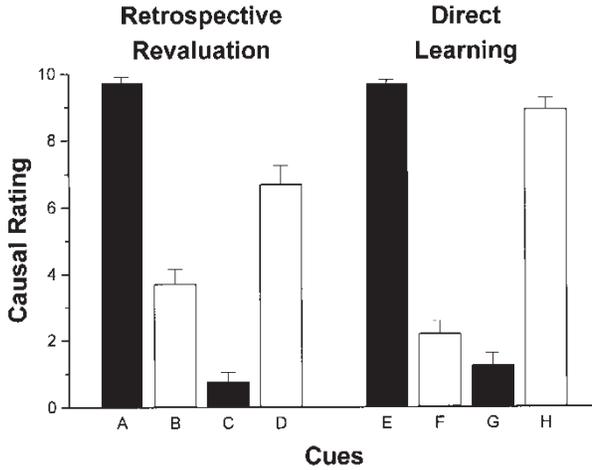


Figure 1. Percentage of trials in each block in which participants predicted an allergic response in Experiment 1; “+” depicts the presence of an allergic reaction, and “-” depicts the absence of an allergic reaction.

on the cue-ratings made before the beginning of the first stage did not find any differences between the various cues,  $F < 1$ . This means that it is unlikely that post-learning differences were due to any specific preconceptions that participants brought to the experiment.

The results for the ratings of the different cues in the retrospective revaluation and in the direct learning conditions can be seen in Figure 2. Ratings of those cues that were presented alone closely reflect their actual contingencies: A and E both received very high ratings and C and G very low ratings. Furthermore, it is also evident that the to-be-blocked target cues B and F were both rated considerably lower than were the release-from-overshadowing cues D and H and that this difference was smaller in the retrospective revaluation condition. This was confirmed statistically by a  $2 \times 2$  repeated-measures ANOVA that included the cues that were not presented alone (B, D, F, and H). This ANOVA compared the factors condition (retrospective revaluation vs. direct learning) and cue (B vs. D and F vs. H). It revealed a significant main effect of cue,  $F(1, 35) = 65.17$ ,  $MSE = 13.02$ ,  $p < .001$ , and more importantly a significant Condition  $\times$  Cue interaction,  $F(1, 35) = 33.45$ ,  $MSE = 3.76$ ,  $p < .001$ . The factor condition did not reach significance,  $F(1, 35) = 1.37$ ,  $MSE = 3.29$ ,  $p > .24$ . To analyse the source of the Condition  $\times$  Cue interaction, we conducted additional paired-samples  $t$ -tests that compared the ratings of the two target cues in each condition separately. For both conditions, these tests revealed that the to-be-blocked target cues were rated significantly lower than were their respective controls,  $t(35) = 3.97$  (retrospective revaluation) and 11.13 (direct learning), respectively, both  $p$ s  $< .001$ . Thus, we found reliable effects of cue selection in both conditions. Since the significant Condition  $\times$  Cue interaction and the results of the  $t$ -tests suggest that there might be a stronger effect in the direct learning condition than in the retrospective



**Figure 2.** Mean causal ratings of the different cues at the end of Experiment 1. Cues that were presented on their own are drawn in black, and cues that were not presented on their own are drawn in white. Error bars depict standard errors of the mean.

revaluation condition, we conducted a final test, in which we compared the amount of cue selection in the two conditions. This test confirmed that the two cues in the former condition were rated more differently than in the latter,  $t(35) = 5.78, p < .001$ . Whether these effects were due to blocking or to release from overshadowing or to both sources of selective learning cannot be determined on the basis of the present data. As this question, however, is not central to the present study, the important observation is that cue selection occurred in both conditions and that the effect was stronger in the direct learning condition than in the retrospective revaluation condition.

Our final and most important analyses concerned the results of the recognition test and evaluated the role of within-compound associations for retrospective revaluation and for direct learning. For each participant we first counted the number of correctly remembered cues for each condition. These served as indicators for the learning of within-compound associations. Participants could get a maximum score of 4 in each condition if they correctly remembered all of the four associates of the blocking and overshadowing cues (e.g., the associates of the respective A and C cues for both replications of the retrospective revaluation condition). On average, they remembered 2.14, 1.83, and 2.00, respectively, of the cues in the retrospective revaluation condition, in the direct learning condition, and for the filler trials. These seem to be comparable to the data of Aitken et al. (2001) who used the same kind of memory test.

Next, we took the ratings of the to-be-blocked cues and their respective control cues and calculated two difference scores. The D – B difference score indicates the degree of retrospective revaluation, and the H – F difference score indicates the degree of cue selection in the direct learning condition. Thus, for each participant we had two memory scores and two difference scores: one for the retrospective revaluation and one the direct learning condition. For each condition we then calculated the product–moment correlation between these two scores to assess whether memory performance was related to the degree of cue selection. Both

correlations were tested one-tailed because predictions of all of the different theoretical accounts were only about values in the positive range (i.e., stronger cue selection effects should be related to better knowledge of within-compound associations). We found a significant correlation only for the retrospective revaluation condition,  $r = .42$ ,  $df = 34$ ,  $p < .01$ . In contrast to this, there was no significant correlation for the direct learning condition,  $r = .11$ ,  $df = 34$ ,  $p > .25$ . When we compared the two correlation coefficients with the test suggested by Steiger (1980), we found that the correlation in the revaluation condition was significantly larger than in the direct learning condition,  $z = 1.65$ ,  $p < .05$  (one-tailed). This means that knowledge of within-compound associations was only meaningfully related to the degree of cue selection in the retrospective revaluation condition and not in the direct learning condition.

From the comparator theory point of view it might be argued that our memory test was not the most appropriate test, as it is the target cue (i.e., the cues that were not presented on their own) that should activate the comparator stimulus. In line with this argument, it might be more appropriate to present the target cues instead of the treatment cues (i.e., the cues from the single cue trials) for the memory test. Yet, as it is the same within-compound association that is assessed in both cases, we do not believe that the kind of cues used in the memory test makes any qualitative difference in the present study.

Taken together, the results of the correlation analyses are clearly at variance with the comparator hypothesis (Denniston et al., 2001; Miller & Matzel, 1988) which predicts that both retrospective revaluation and direct learning should depend on intact within-compound associations. The results are, however, in agreement with both the modified associative theories (Dickinson & Burke, 1996; Van Hamme & Wasserman, 1994) and with the suggestion that rehearsal processes might be responsible for retrospective revaluation (Chapman, 1991). Both of these approaches assume that within-compound associations are only important in the retrospective revaluation condition but should not be crucial in the direct learning condition.

In addition to this, our finding of stronger effects in the direct learning condition than in the retrospective revaluation condition extends results reported by Chapman (1991). In her study, she found evidence for such an asymmetry in the case of forward- versus backward-conditioned inhibition. Additional observations in her paper were also indicative of such an asymmetry for a blocking versus release from overshadowing design, which is now confirmed by our results.

The asymmetry is in line with the findings from the correlation analyses. This is because only participants with knowledge of within-compound associations should be able to process absent cues or rehearse absent trials, while no retrospective revaluation should take place in participants without knowledge of within-compound associations. To the degree to which participants lack such knowledge, this yields a weaker effect in the retrospective revaluation condition. In contrast to this, the occurrence of cue selection is not—or at best only marginally—related to knowledge of within-compound associations in the direct learning condition. A more extended discussion of this aspect is presented in the general discussion.

## EXPERIMENT 2

To investigate whether our findings are generalizable, Experiment 2 adopted the same logic as Experiment 1 but used a backward and forward conditioned inhibition design instead of a

blocking/release from overshadowing design. From an associative point of view, this means that Experiment 2 required learning of inhibitory associations.

The lower half of Table 1 illustrates the design of Experiment 2. For the retrospective reevaluation condition, participants were shown AB– and CD– trials during Stage 1, followed by A+ and C– trials during Stage 2. The A+ trials should retrospectively make B more inhibitory in comparison to D. For the direct learning condition the same basic design was used, with the difference that single cue presentations (E+, G–) now occurred in Stage 1 and compound cue presentations (EF–, GH–) in Stage 2. Finally, additional filler trials ensured that participants also experienced reinforced single-cue and compound trials during both stages and that the numbers of reinforced and nonreinforced trials in each stage were the same. As in Experiment 1, each participant worked on two concurrent replications of the design described in Table 1.

To take the possibility into account that a cue might prevent allergies, we modified the rating scale for Experiment 2. We extended it to negative values, so that it was possible to rate cues differently depending on whether they prevented allergic reactions or just did not cause them.

## Methods

### *Participants*

A different group of 50 University of Marburg students served as participants in Experiment 2. The data of two participants were excluded from the analyses because they did not meet the learning criterion of more than 70% correct predictions during the second half of each stage. Of the remaining participants, 15 were male and 33 were female, and their ages varied between 18 and 38 years, with a median age of 22.5. They received either course credit or were paid for participation. Participants were tested individually and required approximately 30 min to complete the experiment.

### *Apparatus and procedure*

The apparatus, the procedure, and the data scoring and aggregation were identical to Experiment 1 unless stated otherwise.

The instructions from Experiment 1 were modified so that they now covered positive as well as negative contingencies. After the first two paragraphs the instructions now said:

Each food will be presented on the screen by itself. Using a scale from –10 to +10, you should then specify to what degree you believe that this food leads to the occurrence of an allergic response in an ordinary person. +10 means that a food will always cause an allergic reaction, –10 means that a food will always prevent an allergic reaction that may otherwise be caused by another food. 0 means that a food has no influence on the occurrence or the non-occurrence of an allergic response.

The remaining parts of the instructions remained as in Experiment 1. Before Stage 1 and after Stage 2 participants had to rate the causal relationship between each food and the occurrence of an allergic response. Due to the conditioned inhibition design, we changed the rating scale so that it ranged from –10 (*prevents allergic reactions*) to +10 (*causes an allergic reaction*), with 0 (*has no influence on allergic reactions*) as the mid-point. The initial position of the slider was set at 0.

Results and discussion

Figure 3 shows the data for the learning stages. As in Experiment 1, participants' predictions of the allergy outcome closely mirrored the actual contingencies as early as from Block 2 in each stage. This means that participants predicted an allergic response for those cues or cue compounds that were actually reinforced and predicted no allergic response for those that were nonreinforced. Overall, participants correctly predicted the outcome on 96.96 and 94.79% of the trials during the second half of Stage 1 and Stage 2, respectively.

After Stage 2, ratings of the various cues differed significantly,  $F(11, 517) = 85.47, MSE = 10.19, p < .001$ . As there were no differences between ratings before Stage 1,  $F < 1$ , it is unlikely that post-learning differences were due to any pre-experimental biases.

Figure 4 shows the causal ratings of the different cues in the retrospective revaluation and in the direct learning condition. Ratings for those cues that were presented on their own closely reflect their actual contingencies, so that cues A and E received rather high ratings and cues C and G received ratings close to 0. E seems to be rated less positive compared to A, which probably reflects the fact that E was always presented as part of the nonreinforced EF compound during Stage 2, whereas A was always presented on its own and reinforced.

Of greater interest, however, are the ratings of the cues that were not presented on their own. As can be seen in Figure 4, B and F were both rated more negatively than were their respective control cues D and H. Furthermore, this effect seems to be weaker in the retrospective revaluation condition than in the direct learning condition. This was confirmed statistically by a  $2 \times 2$  repeated-measures ANOVA for the cues that were not presented on their own (B, D, F, and H). This ANOVA compared the factors condition (retrospective revaluation vs.

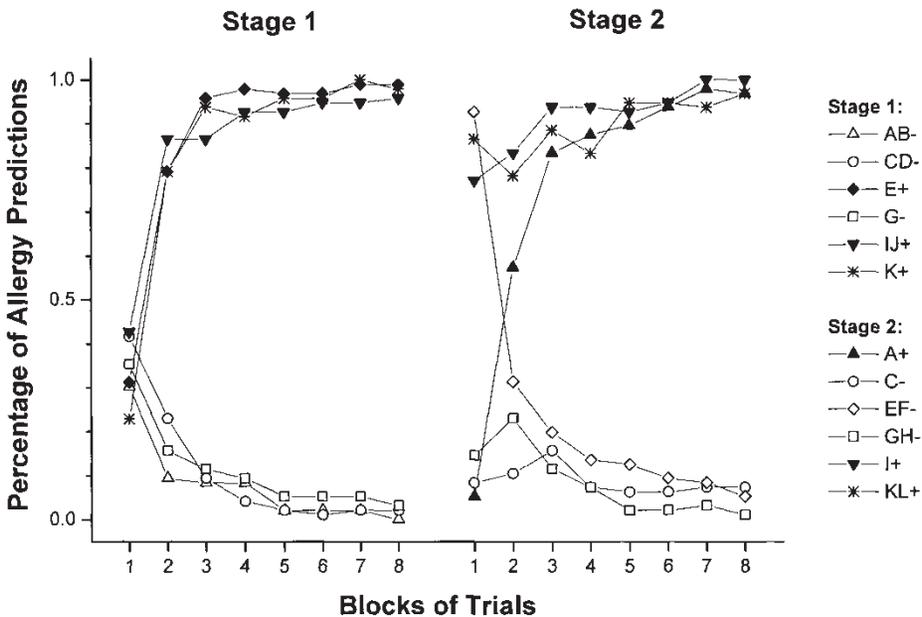
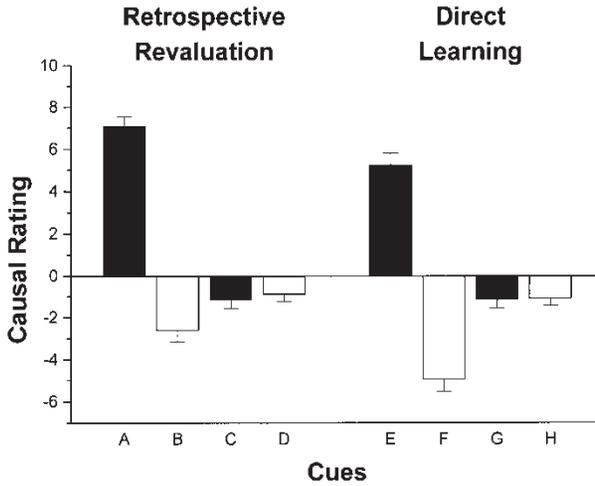


Figure 3. Percentage of trials in each block in which participants predicted an allergic response in Experiment 2; “+” depicts the presence of an allergic reaction, and “-” depicts the absence of an allergic reaction.



**Figure 4.** Mean ratings of the different cues at the end of Experiment 2. Cues that were presented on their own are drawn in black, and cues that were not presented on their own are drawn in white. Error bars depict standard errors of the mean.

direct learning) and cue (to-be-inhibited vs. control). It revealed a significant main effect of cue,  $F(1, 47) = 35.62$ ,  $MSE = 10.50$ ,  $p < .001$ , and, more importantly, a significant Condition  $\times$  Cue interaction,  $F(1, 47) = 9.08$ ,  $MSE = 5.97$ ,  $p < .005$ . Moreover, a significant main effect of condition was found,  $F(1, 47) = 8.77$ ,  $MSE = 8.84$ ,  $p < .006$ , which reflects the more inhibitory ratings in the direct learning condition. Because of the significant Condition  $\times$  Cue interaction, we conducted additional  $t$ -tests to compare the ratings for the two cues in each condition. For both conditions, we found that ratings of the to-be-inhibited target cues (i.e., B and F) were significantly more negative than the ratings of their respective control cues (i.e., D and H),  $t(47) = 3.18$ ,  $p < .004$ , for the retrospective revaluation condition, and  $t(47) = 6.16$ ,  $p < .001$ , for the direct learning condition. Thus, we found reliable conditioned inhibition effects in both conditions. Similar to Experiment 1, however, an additional  $t$ -test confirmed that the effect of the inhibition treatment was stronger in the direct learning condition than in the retrospective revaluation condition, which is indicated by a larger difference of the ratings in the former condition than in the latter,  $t(47) = 3.01$ ,  $p < .005$ .

Again, our final analyses concerned the correlation between the score in the recognition test, which tried to assess knowledge of within-compound associations, and the size of the inhibition effect. On average, participants correctly remembered 1.38, 2.31, and 2.08 of the cues in the retrospective revaluation condition, in the direct learning condition, and for the filler trials, respectively. As in Experiment 1, these memory scores were correlated with the difference scores for the retrospective revaluation condition (D – B) and for the direct learning condition (H – F). The correlation analyses revealed a significant correlation for the retrospective revaluation condition,  $r = .29$ ,  $df = 46$ ,  $p < .03$  (one-tailed), and a nonsignificant correlation for the direct learning condition,  $r = .15$ ,  $df = 46$ ,  $p > .14$  (one-tailed). The difference between these correlations was not significant,  $z = 0.70$ ,  $p > .24$  (one-tailed).

Taken together, the results replicate most of the findings from Experiment 1 and extend them to a conditioned inhibition design. Again, there was no significant correlation between

memory for within-compound associations and the size of the inhibition effect in the direct learning condition, but there was a significant correlation in the retrospective revaluation condition. According to the comparator hypothesis, there should be significant correlations for both conditions to allow for the activation of the appropriate comparator stimuli. Thus, this aspect of our data is again at variance with the comparator account for backward conditioned inhibition. The absence of a significant difference between the correlations for the revaluation condition and the direct learning condition, however, slightly weakens our conclusions. Nevertheless, the fact that the correlation was significant in the retrospective group but not in the direct learning group lends fairly good support to the two accounts, which predict that within-compound associations are only of importance in the retrospective revaluation condition. Further discussion of this point is withheld until the general discussion.

Second, we found that cue selection effects were stronger in the direct learning condition than in the retrospective revaluation condition. This replicates a similar finding by Chapman (1991) and confirms the validity of this result in our Experiment 1. Again, this aspect of our data is at variance with the comparator hypothesis but is in line with the predictions of the two other accounts.

### EXPERIMENT 3

There is recent evidence from several studies that retrospective revaluation is not restricted to cues that have within-compound associations with a cue that is later presented on its own. Instead, retrospective revaluation is also found for cues that have been associated indirectly with such a cue (Blaisdell & Miller, 2001; De Houwer & Beckers, 2002a, 2002b; Denniston, Savastano, Blaisdell, & Miller, 2003). In those experiments, cue A was, for example, presented in compound with another cue B, and cue B was also presented in compound with cue C. Both of these compounds were reinforced (AB+, BC+). Cue C was then presented alone and was either reinforced in one condition or nonreinforced in another condition. Similar to our results in Experiment 1, (first-order) retrospective revaluation of cue B was found after the C- alone trials, so that responding to B was weaker in the condition where C was reinforced than in the condition where C was nonreinforced. But in addition to this, (second-order) retrospective revaluation of cue A was also found with larger responses for A in the group where C was reinforced and smaller responses when C was nonreinforced. As A had no within-compound associations with C, this should not be possible according to the modified associative theories. The explanations that have been offered for those *higher-order revaluation* effects (extended versions of the comparator hypothesis and of cognitive accounts), again crucially depend on knowledge of within-compound associations (Denniston et al., 2001, 2003; De Houwer & Beckers, 2002a, 2002b). According to the extended comparator hypothesis (Denniston et al., 2001), another comparator process moderates the effectiveness of the association between the comparator stimulus and the outcome. For this additional comparator process, the comparator-outcome association is compared to the association of the comparator's comparator stimulus and the outcome. Thus, when A is presented during testing, its potential to elicit responding depends on its associative strength relative to the associative strength of its comparator stimulus B where B's association with the outcome is rendered less effective if B's comparator stimulus C, in turn, has a strong association with the outcome (see Denniston et

al., 2001, for a more detailed discussion). As a consequence, C+ training after prior experience with AB+ and BC+ should make the B–outcome association less effective and thereby increase responding to A. Similarly, C– training should make the B–outcome association more effective and thereby decrease responding to A.

According to the extended comparator hypothesis, two different within-compound associations are needed for the comparator processes: The A–B within-compound association is needed to activate A’s comparator stimulus (i.e., B) and the B–C within-compound association is then needed to activate B’s comparator stimulus (i.e., C). No empirical evidence, however, was presented in any of the studies mentioned above that directly assessed the role of within-compound associations. With an additional experiment we therefore extended our investigation to higher-order cue selection effects and tried to assess the role of within-compound associations for them.

Table 2 shows the design of the experiment. Participants in the retrospective revaluation group received a higher-order retrospective revaluation design, and participants in the direct learning group received a higher-order direct learning design. The higher-order retrospective revaluation design was similar to the designs used by De Houwer and Beckers (2002a, 2002b). In both groups participants experienced AB+, BC+ and DE+, EF+ trials during one stage and C+ and F– trials during the other stage. In the retrospective revaluation group, the various compound trials were presented during the first learning stage and the element trials during the second stage. In the direct learning group, the order of the learning stages was reversed. Additional filler trials ensured that participants also experienced nonreinforced compound trials and that both reinforced and nonreinforced element trials were shown during both stages. In contrast to Experiment 1 and 2 but in line with the experiments by De Houwer and Beckers (2002a, 2002b) participants received only one replication of the respective design.

TABLE 2  
Design of Experiment 3

<i>Group</i>	<i>Type of trials</i>	<i>Stage 1</i>	<i>Stage 2</i>	<i>Memory test</i>
Retrospective revaluation	Second-order compounds	AB+, DE+		A?, D?
	First-order compounds	BC+, EF+		C?, F?
	Single cues		C+, F–	
	Filler	GH–, IJ–, K–, L+	GH–, IJ–, K–, L+	G?, I?
Direct learning	Second-order compounds		AB+, DE+	A?, D?
	First-order compounds		BC+, EF+	C?, F?
	Single cues	C+, F–		
	Filler	GH–, IJ–, K–, L+	GH–, IJ–, K–, L+	G?, I?

*Note:* A to L are different foods, “+” represents the presence of an allergic reaction, and “–” represents the absence of an allergic reaction. Trials were presented eight times per stage. Participants worked on only one replication of the design. “?” means that participants were asked with which other cue the given cue had previously been presented. Before Stage 1 and after Stage 2, participants had to rate all the different cues.

## Methods

### *Participants*

A total of 72 University of Marburg students served as participants, 36 in each group. None of them had taken part in Experiments 1 or 2. An additional participant in the direct learning group was excluded from the analyses because she did not meet the learning criterion of more than 70% correct predictions during the second half of Stage 1 and of Stage 2. Of the participants, 15 were male and 57 were female, and their ages varied between 18 and 38 years, with a median of 23. They either received course credit or were paid for their participation. Participants were tested individually and needed approximately 20 min to complete the experiment.

### *Apparatus and procedure*

The apparatus and the procedure were identical to Experiment 1 unless stated otherwise.

The following foods were used as cues for the experiment: avocados, bananas, cherries, eggs, grapes, kiwi fruits, lemons, mushrooms, nuts, peppers, strawberries, tomatoes. Balancing ensured that each food was used equally often for each kind of cue in Table 2.

The instructions from Experiment 1 were modified to increase comprehension. Furthermore, we introduced additional information concerning the impact of the different cues on the strength of the allergic reactions. In a recent study, De Houwer, Beckers, and Glautier (2002) found that such a manipulation increased the strength of cue selection effects. The instructions now read:

This study is concerned with the question of how people learn about relationships between different events.

Imagine that you are a medical doctor and that one of your patients often suffers from allergic reactions after meals. Your task in this experiment is to find out which kinds of food cause an allergic reaction.

To begin with, however, you should give your current views of different foods. This is to check whether you suspect that certain foods particularly often cause allergies.

Each food will be presented on the screen by itself. Using a scale from 0 to 10 you should then indicate to what degree you suspect that this food leads to the occurrence of an allergic reaction.

This scale will also be shown on the screen.

Give your judgements by moving the slider of the scale with the mouse or by clicking on the slider if you want to retain the default value. After giving a judgement for a food, you can click on the “next” button to move on to the next food.

After participants had rated the different cues they received another window with instructions:

During the course of the experiment you will find out which foods cause allergic reactions in your patient.

You will be shown what your patient has eaten on each day. On some days, he has only consumed a single kind of food, and on some days he has consumed two different foods. Please look at the foods carefully.

Then you will be asked to predict whether or not the patient will show an allergic reaction to the food he consumed. For this prediction, please click on the appropriate response button.

After you have made your prediction, you will be informed whether or not your patient actually showed an allergic reaction. In addition, you will be informed about the strength of the allergic

reaction. A reaction of maximal strength has a value of 20, and a reaction of intermediate strength a value of 10.

Use this feedback to find out to which foods your patient reacts allergically.

Obviously, at first you will have to guess, as you do not know anything about your patient. But eventually you will learn which foods cause allergic reactions in this patient and be able to make correct predictions.

Later, you will be asked to judge again the probability with which the different foods cause allergic reactions in your patient.

For all of your answers, accuracy rather than speed is essential.

Please do not take any notes during the experiment.

If you have got any more questions please ask them now. If you don't have any questions please start the experiment by clicking on the "next" button.

The window for reinforced trials that told participants that their patient showed an allergic reaction was modified in line with De Houwer et al.'s (2002) procedure. On each reinforced trial it now read: "The patient shows an allergic reaction (strength: 10 out of 20)".<sup>4</sup> After the second rating phase, a recognition test was used to assess knowledge of within-compound associations. The cues A, C, D, F, G, and I were presented alone, and participants had to indicate with which other cue they had appeared together as a compound.

### *Data scoring and aggregation*

In contrast to Experiment 1 and 2, neither the trial by trial predictions nor the causal ratings had to be averaged because participants in Experiment 3 only received one replication of the design.

## Results and discussion

Figure 5 depicts the data for the acquisition stages of Experiment 3. The top panel shows the retrospective revaluation group and the bottom panel the direct learning group. In both groups, participants rapidly learned which cues or cue compounds were related to allergic reactions and which were not. Participants in the retrospective revaluation group correctly predicted the outcome on 98.6 and 99.3% of the trials during the second half of Stage 1 and Stage 2, respectively, and participants in the direct learning group on 99.2 and 99.8% of the trials.

After Stage 2, the ratings of the cues were significantly different,  $F(11, 781) = 272.42$ ,  $MSE = 4.37$ ,  $p < .001$ . As in the previous experiments, these differences are unlikely to be due to any pre-experimental biases, because there were no differences in cue ratings before Stage 1,  $F < 1$ .

Figure 6 shows the final ratings for cues A to F after Stage 2. In both groups, the ratings of the cues that were presented alone closely mirror their contingencies, with high ratings for C and low ratings for F. Furthermore, of the first-order cues B was rated lower than E, and of the second-order cues D was rated lower than A in both groups. These effects, however, seemed

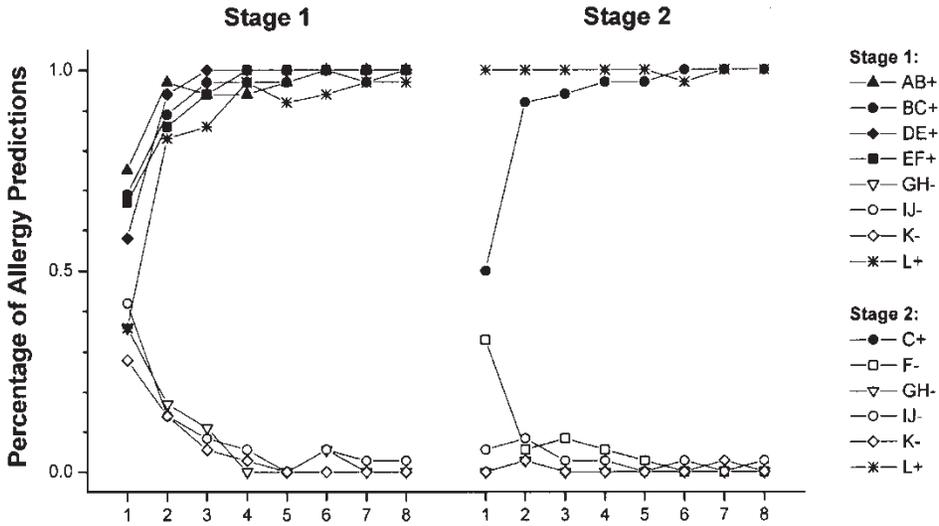
---

<sup>4</sup>In their study, De Houwer et al. (2002) found that this kind of feedback fostered the occurrence of blocking because it demonstrated to participants that the to-be-blocked cue did not have impact over and above the impact of the blocking cue. For the present experiment, however, it seems that this manipulation did not increase the amount of cue competition in comparison to Experiment 1.

to be weaker in the retrospective revaluation group than in the direct learning group. This was confirmed statistically by separate  $2 \times 2$  ANOVAs for the first-order and for the second-order cues. In these ANOVAs, we compared the within-participants factor cue (B vs. E and A vs. D, respectively) and the between-participants factor group (retrospective revaluation vs. direct learning). For the first-order cues B and E, the ANOVA revealed a significant main effect of cue,  $F(1, 70) = 36.24$ ,  $MSE = 10.49$ ,  $p < .001$ , and a significant Cue  $\times$  Group interaction,  $F(1, 70) = 10.51$ ,  $MSE = 10.49$ ,  $p < .003$ . The main effect of group was not significant,  $F(1, 70) = 1.46$ ,  $MSE = 9.18$ ,  $p > .22$ . Separate paired-samples  $t$ -tests confirmed that B was rated lower than E in both groups,  $t(35) = 2.21$ ,  $p < .04$  (retrospective revaluation) and  $t(35) = 5.95$ ,  $p < .001$  (direct learning). A between-group comparison also indicated that the difference between B and E was smaller in the direct learning condition than in the retrospective revaluation condition,  $t_{\text{het}}(67) = 3.24$ ,  $p < .003$ . For the second-order cues A and D, the main effect of cue,  $F(1, 70) = 35.59$ ,  $MSE = 11.62$ ,  $p < .001$ , and of group,  $F(1, 70) = 19.01$ ,  $MSE = 10.07$ ,  $p < .001$ , were significant, as was the Cue  $\times$  Group interaction,  $F(1, 70) = 7.77$ ,  $MSE = 11.62$ ,  $p < .008$ . The significant effect of group reflects the fact that ratings were higher in the direct learning group than in the retrospective revaluation group. For both groups, additional  $t$ -tests confirmed that D was rated lower than A,  $t(35) = 2.12$ ,  $p < .05$  (retrospective revaluation) and  $t(35) = 6.61$ ,  $p < .001$  (direct learning). As for the first-order cues, the difference between the ratings was smaller for the retrospective revaluation group than for the direct learning group,  $t_{\text{hom}}(70) = 2.79$ ,  $p < .008$ . Similarly to Experiment 1 and 2, the significant Cue  $\times$  Group interactions in both ANOVAs were due to the fact that cue selection effects were stronger in the direct learning group than in the retrospective revaluation group.

Our final set of analyses concerned the relationship between knowledge of within-compound associations and the size of the cue selection effects. On average, participants in the retrospective revaluation group correctly remembered 0.89, 1.06, and 1.53 of the first-order compounds, the second-order compounds, and the filler compounds, respectively; participants in the direct learning group remembered 1.03, 1.44, and 1.53 of the respective compounds. Participants could get a maximum score of 2 for each set of cues. To assess the strength of the first-order effects, we calculated the difference between ratings for E and B, and for the second-order effects we calculated the difference between ratings for A and D. The memory scores were then correlated with the size of the cue selection effects. In addition to this, we also calculated correlations between the second-order cue selection effect and joint memory for both the first-order and the second-order within-compound association. This was done because joint knowledge of both within-compound associations is necessary according to the explanations that have been offered for higher-order cue selection (De Houwer & Beckers, 2002a, 2002b; Denniston et al., 2001, 2003). Technically, joint knowledge was operationalized as the product of memory for first-order compounds and memory for second-order compounds. As a consequence, only participants who remembered at least one first-order and one second-order within-compound association could have a value larger than 0. Because several participants remembered either none of the first-order compounds or none of the second-order compounds, calculating this product resulted in a heavily skewed distribution. Therefore, we calculated Spearman rank-order correlations for the present experiment ( $df = 34$  for each correlation). For the retrospective revaluation group, we found a significant correlation between the first-order memory score and first-order retrospective revaluation,  $r = .36$ ,  $p < .02$  (one-tailed), but only a marginally significant correlation between

## Retrospective Reevaluation



## Direct Learning

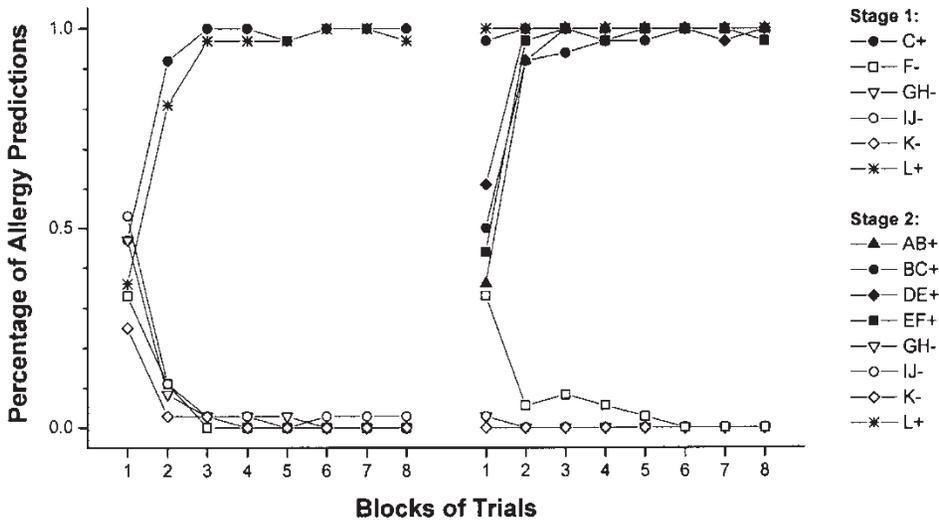
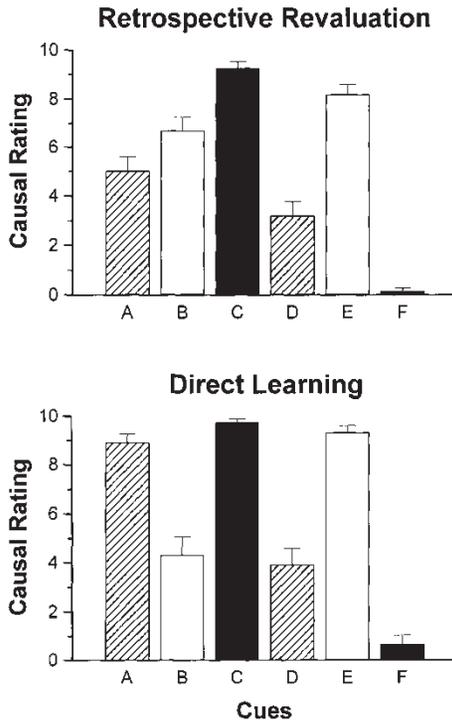


Figure 5. Percentage of trials in each block in which participants predicted an allergic response in Experiment 3 in the retrospective reevaluation group (top panel) and in the direct learning group (bottom panel); “+” depicts the presence of an allergic reaction, and “-” depicts the absence of an allergic reaction.



**Figure 6.** Mean ratings of the different cues at the end of Experiment 3 in the retrospective revaluation group (top panel) and in the direct learning group (bottom panel). Cues that were presented on their own are drawn in black, and cues that were not presented on their own are drawn in white (first-order cues) or are hatched (second-order cues). Error bars depict standard errors of the mean.

the respective memory score and second-order retrospective revaluation,  $r = .27$ ,  $p < .06$  (one-tailed). More important, however, is the fact that there was a large and significant correlation between joint memory for within-compound associations and second-order retrospective revaluation,  $r = .46$ ,  $p < .003$  (one-tailed). In contrast to this, none of the correlations reached significance in the direct learning group,  $r_s = .12$  and  $-.22$ , both  $p_s > .23$  (one-tailed), for the respective memory score and first-order cue selection and second-order cue selection, respectively. Furthermore, the correlation between joint memory for within-compound associations and second-order cue selection was virtually zero,  $r = .01$ ,  $p > .47$  (one-tailed).

When we compared the correlations from both groups statistically, we found that the correlations between memory performance and first-order cue selection did not differ significantly,  $z = 1.05$ ,  $p > .14$ . There were, however, significant differences between the correlations in the two groups for second-order cue selection and memory for second-order within-compound associations,  $z = 2.04$ ,  $p < .03$  (one-tailed), and, more importantly, between second-order cue selection and joint knowledge of within-compound associations,  $z = 2.00$ ,  $p < .03$  (one-tailed).

Taken together, the results from Experiment 3 extend our findings from Experiment 1 and 2 as well as from experiments by others in various regards. First, the retrospective revaluation

group replicated previous findings of higher-order retrospective revaluation (De Houwer & Beckers, 2002a, 2002b; Denniston et al., 2003), and the direct learning group replicated a similar finding of higher-order cue selection effects in rats by Williams (1996; see also Rauhut, McPhee, & Ayres, 1999). Second, we found that both first-order retrospective revaluation and second-order revaluation were weaker than the respective cue selection effects in the direct learning group. Third, our finding of a significant correlation between memory for within-compound associations and (first-order) revaluation in the retrospective revaluation group and of the absence of such a significant correlation between memory and first-order cue selection in the direct learning group replicates the pattern found in Experiment 1. Fourth, our finding that joint knowledge of first-order as well as of second-order within-compound associations is correlated to second-order retrospective revaluation is the first demonstration of such an effect. Although at first glance this finding supports the explanations of retrospective revaluation offered by De Houwer and Beckers (2002a, 2002b) and by Denniston et al. (2003; see also Denniston et al., 2001), the null-correlation in the direct learning group for joint memory and second-order cue selection immediately undermines their explanations. As in Experiments 1 and 2, this is again due to the fact that the comparator process that is assumed to take place during testing should be independent from the order in which the various types of trials were presented. In the general discussion, we discuss to what degree modified associative theories or the rehearsal account can deal with this aspect of our data.

## GENERAL DISCUSSION

There has been a considerable amount of research during the last two decades that has investigated retrospective revaluation; but, there have been surprisingly few studies that evaluated the competing theoretical approaches by comparing conditions that assessed either retrospective revaluation or direct learning. The present experiments were conducted to provide such comparisons.

Experiment 1 used a blocking/release from overshadowing design and Experiment 2 used a conditioned inhibition design. In both experiments, retrospective revaluation effects but not direct learning effects were significantly correlated with knowledge of within-compound associations. Furthermore, cue selection effects were stronger for direct learning conditions than for retrospective learning conditions. Experiment 3 extended these findings to higher order cue selection effects. Most notably, second-order retrospective revaluation was positively correlated with joint knowledge of within-compound associations for both first-order and second-order compounds. This means that participants were more likely to show second-order retrospective revaluation if they knew about both relevant compounds. No such correlation was found for second-order cue selection effects in a direct learning design.

As we have already discussed above, the overall pattern of results clearly undermines the comparator account of retrospective revaluation (Denniston et al., 2001; Miller & Matzel, 1988). This is because the comparator hypothesis predicts significant positive correlations between knowledge of within-compound associations and cue selection effects in both direct learning and retrospective revaluation conditions. In contrast to this, no significant correlations were found in the direct learning conditions. With regard to Experiment 1 and 2, this observation is, however, well in accord with predictions of modified associative theories (Dickinson & Burke, 1996; Van Hamme & Wasserman, 1994) and with the idea that rehearsal

processes might be responsible for retrospective revaluation (Chapman, 1991). Thereby, our results extend findings by Dickinson and his colleagues, who found that manipulations that interfered with learning of the within-compound associations impaired retrospective revaluation but left direct learning intact (Aitken et al., 2001; Dickinson & Burke, 1996; Larkin et al., 1998).

A potential objection with regard to the correlations between memory for within-compound associations and cue selection is that our memory test was not the most appropriate test from the comparator hypothesis point of view. At least for the first-order compounds, we always presented the comparator cue to our participants and asked them to recognize the respective cue with which it was shown together as a compound. This is the activation sequence that is relevant according to the modified associative theories and the rehearsal account. According to the comparator hypothesis, however, the target cue activates its comparator cue at the time of testing so that the sequence in which the representations of the cues are activated is in the opposite direction to the other two accounts. As we have already argued in the discussion of Experiment 1, we see no grounds for supposing that these within-compound associations should be direction-sensitive. We admit, however, that this issue can only be resolved by future research.

With regard to the correlational pattern, it might also be argued that the presence of weak positive correlations in some of the direct learning conditions as well as the absence of a significant difference in some of the comparisons are at variance with our rejection of the comparator account and with our preference for the alternative accounts. First, however, weak positive correlations could be due to the fact that some participants are generally good or generally poor learners. This, in turn, influences both learning of the different contingencies and learning of the within-compound associations. As a consequence, a weak positive correlation might result. Second, as we have argued in the introduction, a weak positive correlation might also be due to the fact that learning about the single cues in the direct learning condition had not reached asymptote at the end of the first learning stage. If Stage 1 trials were rehearsed during the second stage, then this rehearsal would have strengthened the already existing cue selection effect. But even in this case, the correlations in the retrospective revaluation conditions should be greater than the correlations in the direct learning conditions. At least at the descriptive level, this was true in all three experiments, and correlations in the (first-order) retrospective revaluation conditions always exceeded those in the (first-order) direct learning conditions by a factor of two to four. Third, all of the correlations that were predicted by the modified associative theories or by the rehearsal account were significant, whereas none of the correlations that were not predicted by them (or that were only of minor importance according to them) reached significance. Thus, the overall correlational pattern across three experiments seems to be well in accord with these two accounts.

In contrast to Experiment 1 and 2, modified associative theories have considerable difficulties accounting for the results of the retrospective revaluation group in Experiment 3—that is, for higher-order revaluation. The crucial problem is that neither of the second-order cues has any direct within-compound associations with the cues that were presented alone during Stage 2. Therefore, the associative strength of cues A and D should not undergo any change according to the modified associative theories. Denniston et al. (2001) and De Houwer and Beckers (2002a) considered whether a further slight modification of the modified associative theories can extend their explanatory power to higher-order revaluation. If one assumes, for

example, that cue A can be activated indirectly via the B–C and the A–B associations together, then it could also change its associative strength during C+ alone trials. However, this modification makes the predictions of the modified associative theories even worse. This is because it is assumed that changes of the associative strengths of absent cues should be in the opposite direction to changes of present cues. This means that C+ trials should lead to decreases in A's associative strength, while F– trials should lead to increases in D's associative strength (via the E–F and D–E associations). In contrast to this, A was rated more positively than D in the retrospective revaluation condition of Experiment 3.

However, another possibility would be that an absent cue always changes its associative strength in the opposite direction to the cue with which it was presented. Such a possibility could explain the empirical observations because it would predict that, for example, C+ trials lead to decreases in B's associative strength, which, in turn, lead to increases in A's associative strength. Yet, while Van Hamme and Wasserman's (1994) modification of the Rescorla–Wagner model could easily be altered in such a way, it is difficult to conceive how such a modification might be implemented in terms of Dickinson and Burke's modification of Wagner's (1981) SOP model. According to SOP, only cues that are present are able to activate representations of absent cues. These representations, however, should have a lower level of activation than representations that are activated by cues that are actually present. As a consequence of this lower level of activation, they should not be able to activate representations of additional absent cues (Wagner, 1981).

In contrast to the difficulties or the need for additional modifications of associative theories, the rehearsal account can successfully handle higher-order revaluation. One only has to assume that the presentation of C alone not only leads to rehearsal of BC+ trials but that B's within-compound associations with A will also allow rehearsal of AB+ trials. Concurrent rehearsal of both AB+ and BC+ episodes, together with the application of a competitive associative learning rule (e.g., Rescorla & Wagner, 1972), will then make A more excitatory and B less excitatory during C+ trials. Together with rehearsal of the DE+ and EF+ trials during the F– trials, this will result in a difference between the second-order cues A and D in the same way as was observed in the retrospective revaluation condition of Experiment 3.

Another aspect of the present experiments concerns our finding of stronger cue selection effects in the direct learning conditions than in the retrospective revaluation conditions. Cues that were not presented alone were rated more differently in the former conditions than in the latter, which led to significant Cue  $\times$  Stimulus interactions in all experiments. This observation extends Chapman's (1991) findings. Furthermore, similar results seem to be present in animal learning experiments. Matzel, Brown, and Miller (1987, Experiment 1), for example, investigated the effect of repeated presentations of the unconditioned stimulus (US) on later learning about a conditioned stimulus (CS) that was followed by the US. They found that the US preexposure effect (i.e., retardation of developing a conditioned response to the CS after US preexposure) was abolished when the context in which US preexposure had taken place was extinguished either before or after CS+ training. There was, however, a marginally significant effect of stronger recovery from US preexposure when context extinction took place before CS+ training—that is, in the condition that used a direct learning design. Although one might hesitate to give undue weight to this observation, it should be stressed that the authors themselves acknowledged that this was a common observation in similar experiments in their laboratory (Matzel et al., 1987, p. 71).

Denniston et al. (2001) have also recently admitted that asymmetries between direct learning and retrospective revaluation are at variance with the assumption of the comparator hypothesis that trial order (or stage order) should not affect test performance. In addition to this, we have already argued above that the asymmetry in the present experiments is in line with the results from the correlation analyses. For Experiments 1 and 2, this asymmetry is expected by the modified associative theories as well as by the suggestion that rehearsal processes might be responsible for retrospective revaluation. And for Experiment 3 it is expected by the altered version of the Van Hamme and Wasserman (1994) modification of the Rescorla–Wagner theory and by the rehearsal account. Whereas direct learning does not depend on within-compound associations, participants should only be able to retrospectively change associative weights if they know which are the relevant absent cues or the relevant earlier trials. Thus, to the degree that participants lack knowledge of within-compound associations, retrospective revaluation will be impaired, which on average leads to weaker effects of selective learning. Moreover, it has also been suggested that the salience of absent cues might be smaller than those of present cues (Chapman, 1991; Van Hamme & Wasserman, 1994), which would further slow down learning about the former.

Taken together, these two factors—learning about within-compound associations and reduced salience of absent cues—might also be responsible for failures to observe revaluation effects in some experiments (e.g., Grahame, Barnet, & Miller, 1992; Holland, 1999; Kalat & Rozin, 1972; Miller, Hallam, & Grahame, 1990) and for observations that rather extensive revaluation training is often needed in animal learning experiments that yield retrospective revaluation (e.g., Blaisdell, Gunther, & Miller, 1999; Kaufman & Bolles, 1981; Miller, Barnet, & Grahame, 1992).

Given the clear-cut pattern of results in the present experiments, one might wonder why there are nevertheless studies in which no significant differences between cue selection effects in direct learning and retrospective revaluation conditions were found (Aitken, Larkin, & Dickinson, 2000; Dickinson & Burke, 1996; Larkin et al., 1998; Shanks, 1985; Van Hamme, 1994; Wasserman et al., 1996). In several of those studies this might be a question of statistical power, as there often seem to be tendencies toward differences in the same direction as those described in the present paper. In some experiments, however, there does not even seem to be a hint of a difference (Shanks, 1985; Van Hamme, 1994, Experiment 1). From our point of view, the most plausible reason for those findings is that there were comparatively few cues that were used in the training stages of those experiments. In comparison to the present experiments, those studies used designs in which only 2 or 3 cues were used, whereas we, for example, used 24 different cues to set up two replications of the different conditions in Experiments 1 and 2. As a consequence, the demand on working memory was comparatively smaller in the studies in question. This might have made learning the various within-compound associations much less demanding and thus did not impair retrospective revaluation compared to direct learning.

## CONCLUSIONS

On the basis of the present data, the rehearsal account as well as an altered version of the Van Hamme and Wasserman (1994) modification of the Rescorla–Wagner theory are promising theoretical accounts of retrospective revaluation. In contrast to this, several aspects of the

present data are at variance with the comparator theory (Denniston et al., 2001; Miller & Matzel, 1988). Nevertheless, we think that more research is needed that directly compares predictions of the different theoretical approaches. This research might, for example, assess whether the directionality of the within-compound associations is crucial for their correlation with the strength of cue selection effects or whether the difference between the correlations in the direct learning condition versus the retrospective revaluation condition can more reliably be found in larger samples—or in a meta-analysis (Hunter & Schmidt, 1990). Furthermore, we think that future research should also address the question of how great a role rehearsal processes play during associative learning. In the past, researchers have sometimes referred to rehearsal processes as a potential explanation for their results (e.g., Chapman, 1991; Shanks, Darby, & Charles, 1998). However, additional research is needed to assess predictions that may be derived a priori if rehearsal is crucial for the occurrence of retrospective revaluation and other effects.

## REFERENCES

- Aitken, M. R. F., Larkin, M. J. W., & Dickinson, A. (2000). Super-learning of causal judgements. *Quarterly Journal of Experimental Psychology*, *53B*, 59–81.
- Aitken, M. R. F., Larkin, M. J. W., & Dickinson, A. (2001). Re-examination of the role of within-compound associations in the retrospective revaluation of causal judgements. *Quarterly Journal of Experimental Psychology*, *54B*, 27–51.
- Baker, A. G., & Mercier, P. (1989). Attention, retrospective processing and cognitive representations. In S. B. Klein & R. R. Mowrer (Eds.), *Contemporary learning theories: Pavlovian conditioning and the status of traditional learning theory* (pp. 85–116). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Blaisdell, A. P., Gunther, L. M., & Miller, R. R. (1999). Recovery from blocking achieved by extinguishing the blocking CS. *Animal Learning & Behavior*, *27*, 63–76.
- Blaisdell, A. P., & Miller, R. R. (2001). Conditioned inhibition produced by extinction-mediated recovery from the relative stimulus validity effect: A test of acquisition and performance models of empirical retrospective revaluation. *Journal of Experimental Psychology: Animal Behavior Processes*, *27*, 48–58.
- Chapman, G. B. (1991). Trial order affects cue interaction in contingency judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 837–854.
- Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review*, *104*, 367–405.
- Cheng, P. W., & Holyoak, K. J. (1995). Complex adaptive systems as intuitive statisticians: Causality, contingency, and prediction. In H. L. Roitblat & J.-A. Meyer (Eds.), *Comparative approaches to cognitive science* (pp. 271–302). Cambridge, MA: MIT Press.
- De Houwer, J., & Beckers, T. (2002a). Higher-order retrospective revaluation in human causal learning. *Quarterly Journal of Experimental Psychology*, *55B*, 137–151.
- De Houwer, J., & Beckers, T. (2002b). Second-order backward blocking and unovershadowing in human causal learning. *Experimental Psychology*, *49*, 27–33.
- De Houwer, J., Beckers, T., & Glautier, S. (2002). Outcome and cue properties modulate blocking. *Quarterly Journal of Experimental Psychology*, *55A*, 965–985.
- Denniston, J. C., Savastano, H. I., Blaisdell, A. P., & Miller, R. R. (2003). Cue competition as a retrieval deficit. *Learning and Motivation*, *34*, 1–31.
- Denniston, J. C., Savastano, H. I., & Miller, R. R. (2001). The extended comparator hypothesis: Learning by contiguity, responding by relative strength. In R. R. Mowrer & S. B. Klein (Eds.), *Handbook of contemporary learning theories* (pp. 65–117). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Dickinson, A. (2001). Causal learning: An associative analysis (The 28th Bartlett Memorial Lecture). *Quarterly Journal of Experimental Psychology*, *54B*, 3–25.
- Dickinson, A., & Burke, J. (1996). Within-compound associations mediate the retrospective revaluation of causality judgements. *Quarterly Journal of Experimental Psychology*, *49B*, 60–80.

- Dwyer, D. M. (1999). Retrospective revaluation or mediated conditioning? The effect of different reinforcers. *Quarterly Journal of Experimental Psychology*, *52B*, 289–306.
- Grahame, N. J., Barnet, R. C., & Miller, R. R. (1992). Pavlovian inhibition cannot be obtained by posttraining A–US pairings: Further evidence for the empirical asymmetry of the comparator hypothesis. *Bulletin of the Psychonomic Society*, *30*, 399–402.
- Holland, P. C. (1999). Overshadowing and blocking as acquisition deficits: No recovery after extinction of overshadowing or blocking cues. *Quarterly Journal of Experimental Psychology*, *52B*, 307–333.
- Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis*. Newbury Park, CA: Sage.
- Kalat, J. W., & Rozin, P. (1972). You can lead a rat to poison but you can't make him think. In M. E. P. Seligman & J. L. Hager (Eds.), *Biological boundaries of learning* (pp. 115–122). New York: Appleton–Century–Crofts.
- Kamin, L. J. (1968). “Attention-like” processes in classical conditioning. In M. R. Jones (Ed.), *Miami Symposium on the Prediction of Behavior, 1967: Aversive stimulation* (pp. 9–31). Coral Gables, FL: University of Miami Press.
- Kaufman, M. A., & Bolles, R. C. (1981). A nonassociative aspect of overshadowing. *Bulletin of the Psychonomic Society*, *18*, 318–320.
- Larkin, M. J. W., Aitken, M. R. F., & Dickinson, A. (1998). Retrospective revaluation of causal judgments under positive and negative contingencies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 1331–1352.
- Mackintosh, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review*, *82*, 276–298.
- Matzel, L. D., Brown, A. M., & Miller, R. R. (1987). Associative effects of US preexposure: Modulation of conditioned responding by an excitatory training context. *Journal of Experimental Psychology: Animal Behavior Processes*, *13*, 65–72.
- Matzel, L. D., Schachtman, T. R., & Miller, R. R. (1985). Recovery of an overshadowed association achieved by extinction of the overshadowing stimulus. *Learning and Motivation*, *16*, 398–412.
- Miller, R. R., Barnet, R. C., & Grahame, N. J. (1992). Responding to a conditioned stimulus depends on the current associative status of other cues present during training of that specific stimulus. *Journal of Experimental Psychology: Animal Behavior Processes*, *18*, 251–264.
- Miller, R. R., Hallam, S. C., & Grahame, N. J. (1990). Inflation of comparator stimuli following CS training. *Animal Learning & Behavior*, *18*, 434–443.
- Miller, R. R., & Matute, H. (1996). Biological significance in forward and backward blocking: Resolution of a discrepancy between animal conditioning and human causal judgment. *Journal of Experimental Psychology: General*, *125*, 370–386.
- 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.
- Pearce, J. M. (1994). Similarity and discrimination: A selective review and a connectionist model. *Psychological Review*, *101*, 587–607.
- Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*, *87*, 532–552.
- Rauhut, A. S., McPhee, J. E., & Ayres, J. J. B. (1999). Blocked and overshadowed stimuli are weakened in their ability to serve as blockers and second-order reinforcers in Pavlovian fear conditioning. *Journal of Experimental Psychology: Animal Behavior Processes*, *25*, 45–67.
- Rescorla, R. A., & Cunningham, C. L. (1978). Within-compound flavor associations. *Journal of Experimental Psychology: Animal Behavior Processes*, *4*, 267–275.
- Rescorla, R. A., & Durlach, P. J. (1981). Within-event learning in Pavlovian conditioning. In N. E. Spear & R. R. Miller (Eds.), *Information processing in animals: Memory mechanisms* (pp. 81–111). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- 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 theory and research* (pp. 64–99). New York: Appleton–Century–Crofts.
- Shanks, D. R. (1985). Forward and backward blocking in human contingency judgement. *Quarterly Journal of Experimental Psychology*, *37B*, 1–21.
- Shanks, D. R., Darby, R. J., & Charles, D. (1998). Resistance to interference in human associative learning: Evidence of configural processing. *Journal of Experimental Psychology: Animal Behavior Processes*, *24*, 136–150.

- Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, 87, 245–251.
- Van Hamme, L. J. (1994). *Associative and statistical accounts of cue competition in causality judgments*. Unpublished doctoral dissertation, University of Iowa, Iowa, USA.
- Van Hamme, L. J., & Wasserman, E. A. (1994). Cue competition in causality judgments: The role of nonpresentation of compound stimulus elements. *Learning and Motivation*, 25, 127–151.
- Wagner, A. R. (1981). SOP: A model of automatic memory processing in animal behavior. In N. E. Spear & R. R. Miller (Eds.), *Information processing in animals: Memory mechanisms* (pp. 5–47). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Wasserman, E. A., & Berglan, L. R. (1998). Backward blocking and recovery from overshadowing in human causal judgement: The role of within-compound associations. *Quarterly Journal of Experimental Psychology*, 51B, 121–138.
- Wasserman, E. A., Kao, S.-F., Van Hamme, L. J., Katagiri, 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, pp. 207–264). San Diego, CA: Academic Press.
- Williams, B. A. (1996). Evidence that blocking is due to associative deficit: Blocking history affects the degree of subsequent associative competition. *Psychonomic Bulletin & Review*, 3, 71–74.

*Original manuscript received 2 January 2002*

*Accepted revision received 3 February 2003*