

The 28th Bartlett Memorial Lecture

Causal learning: An associative analysis

Anthony Dickinson

University of Cambridge, Cambridge, U.K.

The concordance between performance and judgements of the causal effectiveness of an instrumental action suggests that such actions are mediated by causal knowledge. Although causal learning exhibits many associative phenomena—blocking, inhibitory or preventative learning, and super-learning—judgements of the causal status of a cue can be changed retrospectively as a result of learning episodes that do not directly involve the cue. In order to explain retrospective reevaluation, a modified associative theory is described in which the learning processes for retrieved cue representations are the opposite to those for presented cues, and this theory is evaluated by studies of the role of within-compound associations in retrospective reevaluation and blocking. However, this modified theory only applies when the within-compound association represents a contiguous rather than a causal cue relationship.

Causal learning and representation is a fundamental form of cognition, if not the fundamental form. Without the capacity to learn about and represent the causal relationships between our actions and their consequences, the mind would be radically disconnected from the world. However detailed and rich our knowledge, however sophisticated and complex our inferences and planning, cognition would be impotent if our thoughts could not be expressed in goal-directed action that fulfils our needs and desires. And, for the behavioural expression of cognition to be effective, it must be grounded upon knowledge of the causal relationships between our actions and their consequences in the world. So causal action–outcome knowledge, through its interaction with motivational processes, is the final common pathway for the expression of cognition in action, and it is for this reason that causal learning can be viewed as a primary, or fundamental form of cognition.

It was this line of reasoning that led David Shanks and myself to embark, nearly 20 years ago, on an investigation of the relationship between causal cognition and instrumental action (Dickinson, Shanks, & Evenden, 1984). At that time, the study of instrumental action had been largely conducted within the animal conditioning laboratory, and apart from the

Requests for reprints should be sent to A. Dickinson, Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge, CB2 3EB, U.K. Email: ad15@hermes.cam.ac.uk

This paper is based upon the 28th Sir Frederic Bartlett Lecture, delivered at the July 2000 EPS meeting at Cambridge University. The research reported in this paper was funded by BBSRC grants. I should like to thank David Shanks, Michael Aitken, and Nicholas Mackintosh for their comments on an earlier draft of this lecture and Helen Brough, Helen Corbett, and Jonathan Lee for their assistance in conducting some of the experiments.

pioneering work of Jenkins and Ward (1965) and Alloy and Abramson (1979), the processes by which humans learn about the causal consequences of their actions and the relationship between this learning and instrumental performance had received little or no attention. It is true that students of conditioning had long been concerned with the issue of whether or not instrumental conditioning can occur in the absence of “awareness” of the action–outcome contingency (see Brewer, 1974), but this research was not concerned with the processes of causal learning. So one of the first issues that we addressed is whether common learning processes mediate instrumental performance and the acquisition of causal knowledge.

Contiguity and contingency

The strategy we adopted was that of investigating the concordance between causal judgements and instrumental performance across variations in the main parameters of an action–outcome relationship. The argument was that such concordance favours the claim that instrumental action is based upon causal cognition. It is, of course, a trivial matter to show, by demonstration alone, that humans act on the basis of explicit causal information; at issue, however, is whether instrumental behaviour acquired through experience with an action–outcome relationship is also based upon causal knowledge.

Studies of animal conditioning have identified two main parameters that impact on instrumental performance. The first is the temporal contiguity between the action and its outcome. If a lever press by a rat delivers a food reward with some fixed probability, the rate of lever pressing declines systematically when a delay of a few seconds is introduced between each press and the reward (e.g., Dickinson, Watt, & Griffiths, 1992). In order to investigate whether the same is true for humans, we adopted an equally impoverished and simple task¹ (Shanks & Dickinson, 1991). The instrumental action was a press on the space bar of a keyboard and the outcome the illumination of a figure on the computer monitor, and the subjects were simply asked to maximize the total number of points gained, with each action costing 1 point and each outcome worth 3 points. As, on average, 9 out of 10 presses generated an outcome, the optimal strategy was to press as frequently as possible. Whatever the delay between the action and outcome, 10 presses would have gained, on average, 27 points, at the cost of only 10 points.

Our subjects did not always adopt this strategy, however. Although they pressed the space bar at a high rate when the illumination of the figure followed immediately, imposing a delay of only 2 s between the response and the outcome produced a precipitous decline in the rate of responding. In fact, with 2- and 4-s action–outcome delays, the rate of responding was not reliably different from a non-contingent, control condition. In this non-contingent condition, we simply recorded the temporal pattern of outcomes generated by the subject in the matched contingent condition and replayed it independently of responding. Therefore, responding on this non-contingent schedule has no effect on the likelihood of an outcome.

¹ Our original instrumental task (Dickinson et al., 1984) used a richer scenario. As is so often the case, however, Wasserman and his colleagues independently embarked on a similar research programme on instrumental causal learning at about the same time (Chatlosh, Neunaber, & Wasserman, 1985), and for these studies we adopted their simpler procedure because it is more akin to those used in standard animal operant conditioning.

So humans, like rats, are very sensitive to action–outcome contiguity. The critical issue, however, is whether beliefs about the strength of the causal relationship between action and outcome are equally sensitive to temporal contiguity. To address this issue, we tested a second group of subjects under the same conditions, except that their task was not to maximize their points total but, rather, to judge the strength of the causal relationship after equivalent exposure to action–outcome contingency. Causal judgements also declined with outcome delay until, with a 4-s delay, the causal relationship between action and outcome was judged to be about half as strong as with no delay.

This concordance between judgements and responding accords with the claim that causal beliefs play a generative role in instrumental performance. It is true that the judgements discriminated between the contingent and non-contingent conditions under the 2-s and 4-s delays, whereas the performance did not. In evaluating this difference, however, it is important to note that the payoff structure of this task is such that it is only worth responding if a press causes an outcome with a probability greater than .3. So, even though a subject might well believe that an action has a positive causal efficacy, the failure to respond on the basis of the belief can be a rational choice. It might also be thought surprising that our subjects did not detect the action–outcome relationship across such short delays. However, we gave only a brief, 2-min training period under each delay, and I am confident that with more training they would have outperformed a rat's ability to detect the contingency between a lever press and food with a delay of over a minute (Dickinson, Watt, & Varga, 1996).

The second main parameter of animal instrumental conditioning is the strength of the contingency between the action and outcome. I have already illustrated the contrast between the contingent and non-contingent schedules in our delay study. In this case, however, the effect of contingency can simply be attributed to variations in the contiguity between action and outcome. The outcomes were necessarily paired with the presses in the contingent condition with the 0-s delay, but were only fortuitously, and therefore infrequently, paired in the matched non-contingent condition. However, the action–outcome contingency can be varied while keeping the probability of a paired outcome constant by altering the likelihood that the outcome occurs in the absence of an action or, in other words, by varying the probability of an unpaired outcome. When the probability of the action–paired outcome is greater than that of the unpaired outcome, more outcomes can be gained by responding than by not responding, and there is a causal relationship between action and outcome. By contrast, under a non-contingent schedule in which the probability of the paired and unpaired outcomes are the same, just as many outcomes occur when the agents never respond as when they respond at every opportunity.

Hammond (1980) was the first to demonstrate that rats are sensitive to variation in the strength of the action–outcome contingency brought about by changes in the probability of an unpaired outcome, and the same is true of human subjects. With the probability of the paired outcome again set at .9, we (Shanks & Dickinson, 1991—see also Chatlosh et al., 1985) found that the response rate and the causal judgements declined in a similar fashion as the probability of unpaired outcomes was raised.

In summary, we observed a good concordance between instrumental performance and causal judgements across variations in both action–outcome contiguity and contingency. At the very least, this concordance provides correlational evidence that instrumental action is based upon an acquired belief about the causal efficacy of the action.

Associative and causal learning

Within the instrumental conditioning literature, sensitivity to contiguity and contingency is usually explained in terms of associative learning theory according to which the rate of an instrumental response is assumed to reflect the strength of the association between representations of the action and the outcome acquired through pairings of these two events (e.g., Dickinson, 1994; Rescorla, 1991). The concordance between instrumental performance and causal judgement suggested to us (Dickinson et al., 1984; Shanks & Dickinson, 1988) that causal learning may also be associative with judgements reflecting the strength of the action–outcome association.

At the time, the main evidence in favour of an associative account came from studies of the acquisition of causal judgements and their terminal values under variations in action–outcome contingencies (e.g., Dickinson et al., 1984; Shanks, 1987; Shanks & Dickinson, 1988; Shanks, Lopez, Darby, & Dickinson, 1996; Wasserman, Elek, Chatlosh, & Baker, 1993). There is, however, a major problem with investigating learning using an instrumental procedure in that this procedure confounds the independent and dependent variables. The occurrence of the putative cause, along with the outcome, are the major independent variables that can be manipulated in the study of causal learning, and yet an instrumental procedure also measures learning by the frequency of the cause—that is, the action. This confound led to a shift away from instrumental to Pavlovian procedures for the analysis of learning in the conditioning laboratory, and, more recently, we have also adopted a Pavlovian-type procedure to study causal learning.

In this procedure, the participant plays the role of a food allergist whose task is to judge the extent to which different foods cause an allergic reaction in a hypothetical patient. These judgements are based upon the information provided by a number of training trials in each of which the participant is presented with the food(s) eaten in a meal and is then required to predict whether or not this meal will cause an allergic reaction. Immediately after the predictive response, the participant is informed about whether or not an allergic reaction occurs following that meal. At each stage of training, the subjects are given sufficient training to allow them to learn to predict the outcome reliably before they are finally asked to rate how effective each food is in causing the allergic reaction.

Two further features of our procedure should also be noted. First, our studies typically employed within-subject designs so that each subject was trained concurrently on a number of different contingencies between various foods and the allergic reaction. These contingencies were carefully chosen to be diagnostic of associative learning. Second, more than one food acted as each type of cue, so that the participants were required to learn not only about a number of contingencies concurrently, but also about the causal status of more than one food within each contingency. In describing our experiments, however, I shall refer to the foods that bear the same relationship to the outcome by a single, common letter.

Although associative theories of conditioning come in many forms, most of them, at their heart, claim that only surprising or unexpected outcomes support sustained learning. For some theorists (e.g., McLaren, Kaye, & Mackintosh, 1989; Pearce, 1987; Sutton & Barto, 1981; Rescorla & Wagner, 1972; Wagner, 1981), the occurrence of a surprising outcome has a direct impact on the associative strength of a cue, whereas for others surprise operates, at least in part, by modulating the attentional salience or associability of the cue (e.g., Pearce

& Hall, 1980; Mackintosh, 1975). Whatever the specific process involved, however, the role of surprise in learning is most simply demonstrated by Kamin's (1969) blocking effect.

Blocking

In the first stage of our blocking procedure (Aitken, Larkin, & Dickinson, 2000), a single food X was paired with the outcome (see Figure 1). As this outcome was initially unpredicted or surprising, it generated a food-X-outcome association that strengthened across trials until the occurrence of the outcome was fully predicted by food X. Then, in Stage 2, a compound meal consisting of two foods—the pretrained food X and a novel food B—was paired with the outcome. However, as the occurrence of the outcome following these compound meals was fully predicted by food X and therefore not surprising, little or no association should have been formed between the added food B and the outcome. In agreement with this prediction, the final causal ratings for food B were very low and significantly less than those for a control food C (see Figure 1). Food C received exactly the same training as food B except for the fact that the other food in the compound meal, food Y, had not been pretrained in the first stage. Thus, the pretraining of food X rendered the allergic reaction in the second stage unsurprising and “blocked” learning about food B. The blocking of causal and predictive learning is a well-established phenomenon in human learning that has been demonstrated in a variety of scenarios (e.g., Chapman & Robbins, 1990; Dickinson et al., 1984).

Preventative learning

According to associative theory, not only does pairing a cue with a surprising outcome support learning, but so does the pairing of a cue with the unexpected omission of a predicted outcome. In this case, however, the pairing results in the formation of an inhibitory rather than excitatory association between the cue and outcome representations. The conditions for the formation of such an inhibitory association are illustrated in the preventative contingency of Figure 1. In addition to the meals of the blocking contingency, the patient also received another type of compound meal during Stage 2 in which a further food P was eaten with the pretrained food X, but in this case the allergic reaction did not occur. In this contingency, therefore, food X was paired with the surprising omission of the outcome, which was expected because of the presence of food X in the compound meal. In other words, food P acted as a prophylactic for the allergic reaction caused by food X.

Within the context of causal learning, an inhibitory association represents a *preventative* relationship between a putative cause and an outcome, in contrast to the *generative* causal relationship represented by an excitatory association. In order to assess whether our participants learned about the preventative relationship, we (Aitken et al., 2000) asked them also to rate the causal effectiveness of food P under the instructions that preventative causes should be assigned negative ratings. As Figure 1 illustrates, on average cue P attracted highly negative ratings, thereby replicating previous demonstrations of preventative learning in predictive and causal scenarios (e.g., Chapman, 1991; Williams, 1995).

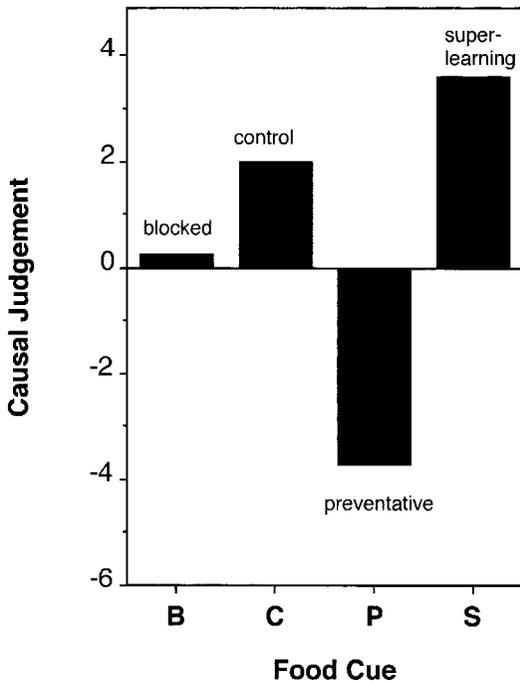
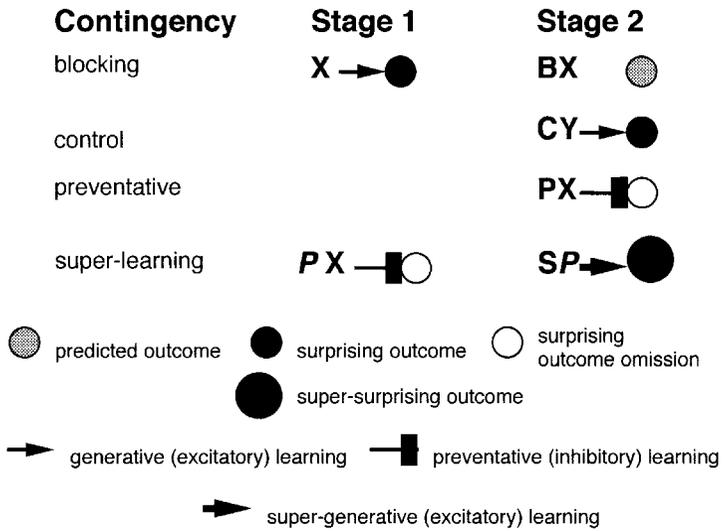


Figure 1. The food-cue–outcome pairings and the standard associative learning processes for the blocking, control, preventative, and super-learning contingencies and the mean causal judgements of the B, C, P, and S food cues from these contingencies (Aitken et al., 2000).

Super-learning

The occurrence of preventative learning allowed us to test one further prediction of associative theory. As I have already noted, generative learning occurs whenever a cue is paired with an unexpected outcome, with the amount of learning depending upon the extent to which the outcome is surprising. Therefore, most surprising outcomes should support greater generative learning than do less surprising ones. The super-learning contingency in Figure 1 illustrates a procedure for rendering the outcome super-surprising (Rescorla, 1971).

In the first stage, another food *P* was established as a preventative cause of the allergic reaction by intermixing meals in which food *X* alone was paired with the outcome with other, compound meals in which foods *X* and *P* were presented without the outcome. Then, in the second stage, the participants also received another type of compound meal consisting of food *P* and a further, novel food *S*, and this meal was consistently followed by the allergic reaction for a number of trials. The important feature of this contingency is that the occurrence of the outcome on the compound trials should have been super-surprising. Stage 1 established food *P* as a preventative cause, so that the participants should have anticipated that the presence of this food in the compound meals of Stage 2 should have prevented the allergic reaction. Therefore, the fact that the reaction did occur would have been very unexpected and should have supported extra-strong generative learning to food *S*. In agreement with this prediction, food *S* attracted the highest causal ratings of all the foods and significantly more than the control food *C*, which was trained in compound with a novel cue.

In summary, this study (Aitken et al., 2000) demonstrated that human causal learning manifests a number of effects predicted by standard associative theories—blocking, preventative learning, and super-learning—and encouraged us to take such theories seriously as accounts of causal learning.

Retrospective reevaluation

Although associative theories explain many of the interactions between potential causes, the standard versions of these theories impose major restrictions on causal inference. This point can be illustrated by considering a contingency in which the two stages of the blocking contingency (see Figure 1) are reversed to yield the retrospective blocking contingency illustrated in Figure 2. Under this contingency, the *BX* compound meal is paired with the outcome in the first stage and food *X* alone with the outcome in the second stage. At the time of causal judgement, the information available to the participants about the contingency between food *B* and the allergic reaction is identical in the forward (Figure 1) and retrospective cases (Figure 2). In both cases, they have observed that the allergic reaction occurs following meals both with and without food *B*. And yet standard associative theories predict very different judgements about the causal effectiveness of cue *B* in the two cases.

To recap, associative theories predict little or no causal learning to cue *B* in the forward blocking contingency because the outcome paired with *B* is predicted by the pretrained food *X* (see Figure 1). By contrast, food *B* should attract relatively high causal ratings after training on the retrospective blocking contingency. This food is paired with a surprising outcome in the first stage and should therefore gain as much associative strength as the control food *C*,

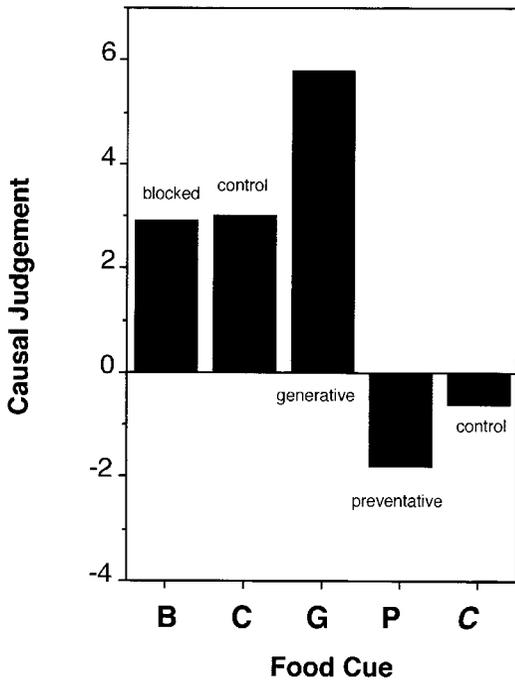
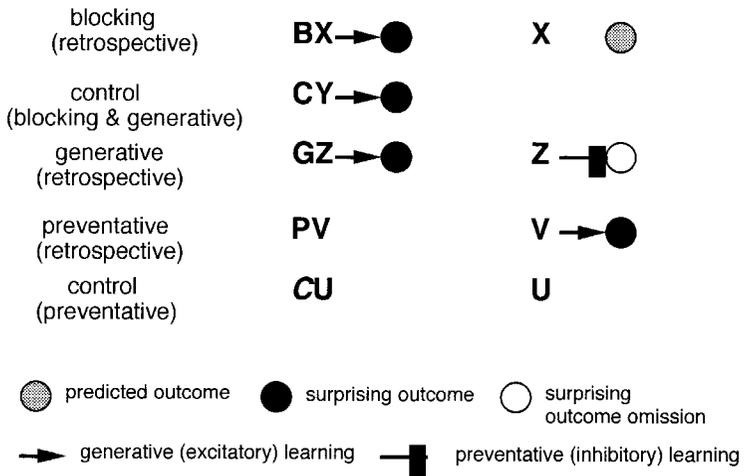


Figure 2. The food-cue-outcome pairings and the standard associative learning processes for the blocking, generative, and preventative retrospective contingencies and their respective control contingencies and the mean causal judgements of the B, C, G, P, and C food cues from these contingencies (Larkin et al., 1998).

which receives the equivalent compound training in the first stage (see Figure 2). Although food X is subsequently paired with a predicted outcome in the second stage,² this training should have no impact on the associative strength of food B, because this food is not present in the meals presented in the second stage. It is a fundamental assumption of standard associative theories that learning about a cue can only occur on episodes in which the cue is present. In agreement with this prediction, Figure 2 shows that we (Larkin, Aitken, & Dickinson, 1998) did not observe retrospective or backward blocking.³ In fact, the causal ratings for food B were indistinguishable from those for the control cue, food C.

Although the absence of retrospective blocking accords with associative theory, we did observe another, generative form of retrospective revaluation in these studies (Larkin et al., 1998). The contingency producing this revaluation is also illustrated in Figure 2. Again, we paired a compound meal, consisting of foods G and Z, with the outcome in the first stage, but in this case we presented food Z alone in the absence of the outcome in the second stage. A rational analysis of this contingency might well support a revaluation of the causal status of cue G. The presentation of food Z alone in the second stage demonstrates that this food does not cause an allergic reaction, thereby resolving any ambiguity about which food caused the outcome in the first stage in favour of food G—a resolution that might well elevate its causal status. Figure 2 shows that just such an elevation occurred in that the causal judgements for food G were almost twice those for the control food C and the blocked food B.⁴

This form of retrospective revaluation challenges standard associative theory in that generative retrospective revaluation involves a change in the causal status of a cue as a result of training episodes that do not include that cue. It is clear, therefore, that the theory has to be modified and extended if it is to be retained as a plausible account of causal learning.

Learning about absent causes

A clue to how the theory might be modified can be gleaned from a cartoon that appeared in the *Christian Science Monitor* under the caption “Russian army cannot pay its soldiers” (J. Kruschke, personal communication). The cartoon portrays two Russian soldiers on sentry duty in some bleak Siberian posting, with one musing to the other: “I’m beginning to *associate* Communism with paychecks. That’s not right, is it?” (italics mine). In the present context, what this episode illustrates is the formation of a generative association between two absent entities, Communism and a paycheque, which is exactly the association required to explain generative retrospective revaluation. In fact, this episode is an example of the retrospective revaluation of a causal relationship in which Communism is the cause and the paycheque the outcome. An associative analysis of this episode could take the following form. The cartoon suggests that the soldiers are veterans of the Red Army, so that past service led to them to associate both Communism and paycheques with previous postings. As a consequence, the current

² For the sake of simplicity, this analysis ignores overshadowing between foods B and X during Stage 1, which would render the outcome under-predicted by food X in Stage 2.

³ Other studies (e.g., Shanks, 1985; Wasserman & Berglan, 1998) have reported retrospective or backwards blocking. Larkin et al. (1998) offer an explanation of why the magnitude of retrospective blocking may vary.

⁴ An analogous effect, unovershadowing, is documented in the animal learning literature (e.g., Kaufman & Bolles, 1981; Miller, Barnet, & Grahame, 1992—but see Holland, 1999).

posting retrieves memories of both Communism and paycheques, and the conjoint activation of the representations of these events in memory is sufficient, even in their absence, to produce an association between them.

The application of this idea to the retrospective contingencies is illustrated in Figure 3.⁵ The assumption is that during training with the compound meals of Stage 1 associations are formed not only between foods and the outcome but also between the foods themselves—foods B and X in the case of the blocking contingency and foods G and Z in the case of the generative contingency. These within-compound associations represent the fact that these two foods occurred together in the meals. As a consequence, when either food X or food Z is presented alone during Stage 2, it activates not only a representation of the outcome but also a representation of the other food in the meal—either food B or food G, respectively.

Associative retrieval of the representation of food B in the second stage of the retrospective blocking contingency should have no effect on the causal status of this food, as indeed it did not (see Figure 2). To recap, the allergic reactions that occurred during Stage 2 were predicted by food X and should therefore not have engaged associative learning. By contrast, the modified theory predicts reevaluation for the retrospective generative contingency. Once again, presentations of food Z during the second stage should have associatively activated representations of both food G and the allergic reaction. However, the fact that this predicted outcome surprisingly did not occur is, according to the modified theory, the condition for excitatory or generative associative learning between an associatively activated cue representation, such as food G, and the associatively activated outcome representation. Consequently, in line with the empirical evidence (see Figure 2), the causal status of food G should have been enhanced retrospectively.

According to this modified theory, the type of learning supported by an associatively activated or retrieved cue representation is the opposite of that for a presented cue. Whereas outcome omission supports inhibitory or preventative learning when paired with food Z, it produces excitatory or generative learning when paired with the retrieved representation of food G (see Figure 3). This claim immediately raises the issue of whether this reversal in the form of learning is a general property of associative processes. If it is, it should be possible to establish a cue as a preventative cause by pairing its representation with the surprising presentation of the outcome. In other words, a general version of the modified theory predicts that preventative causal status can be acquired retrospectively.

Figure 3 illustrates the predictions of the modified theory for a retrospective preventative contingency. In the first stage, the participants receive a compound meal composed of foods P and V that do not cause any allergic reaction, so that all that should have been learned is a within-compound association between the two foods. As a consequence, the presentation of food V alone in Stage 2 should activate a representation of food P via the within-compound association. The occurrence of the unpredicted outcome on these trials, while supporting excitatory or generative learning to food V, should form an inhibitory or preventative association between food P and the outcome. Figure 2 shows that this prediction of the modified theory was upheld (Larkin et al., 1998a). Although the absolute magnitude was not large, our subjects rated food P more negatively than a control food C that

⁵ Van Hamme and Wassermann (1994), Dickinson and Burke (1996), Larkin et al. (1998) and Graham (1999) have all presented modifications of standard associative theories to account for retrospective reevaluation.

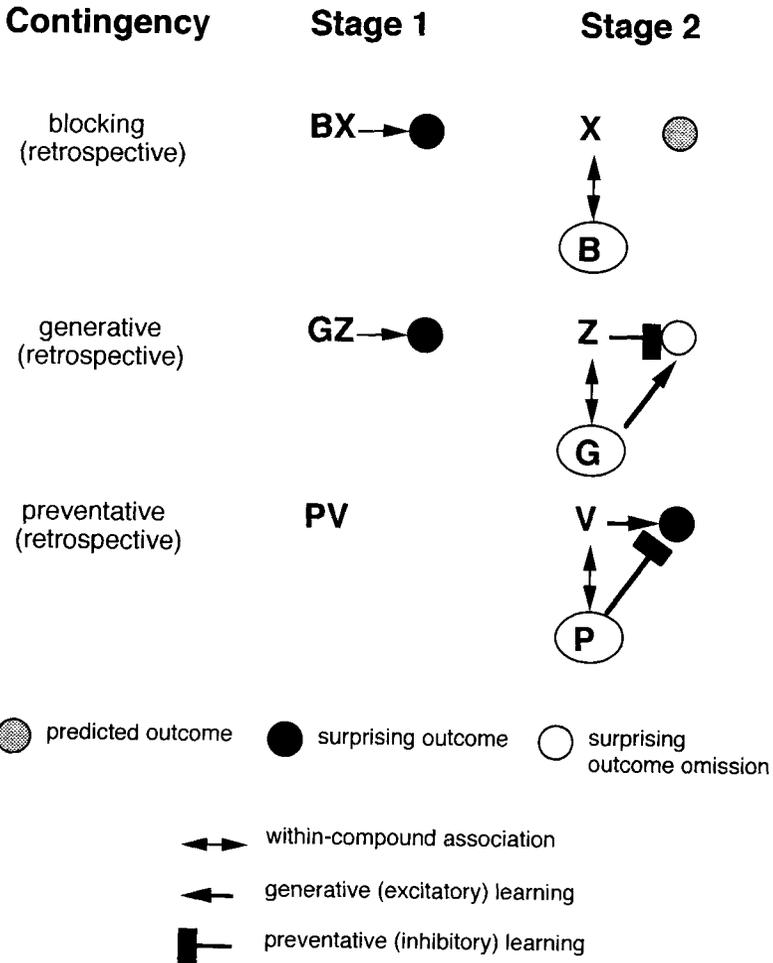


Figure 3. The food–cue–outcome pairings, the standard associative processes for cue presentations, and the modified associative learning processes for cue representations activated by within-compound associations in the blocking, generative, and preventative retrospective contingencies.

was trained under a similar contingency, except that the outcome was not presented during the second stage. Again, the reliability of retrospective preventative learning has been established in a variety of procedures (e.g., Chapman, 1991; Williams & Docking, 1995).

Role of within-compound associations

The process at the heart of the associative account of retrospective revaluation is the retrieval or activation of a cue representation via a within-compound association. It is this process that allows for a change in the casual status of a cue even in the absence of the cue itself. In order to assess the role of within-compound associations, we have investigated the effect of manipulations designed to minimize within-compound learning (Dickinson &

Burke, 1996; Larkin et al., 1998). Most recently, we (Aitken, Larkin, & Dickinson, 2001) have used a concurrent interference task designed to reduce the strength of the association between the foods constituting a compound meal. Our argument was that any treatment that interferes with the formation of within-compound associations should equally interfere with the retrieval of the representation of the absent cue during the second stage of training and thereby reduce the magnitude of retrospective reevaluation.

We chose mental arithmetic as the interfering task largely on empirical grounds, although it should be noted that there is good evidence that mental arithmetic makes demands on executive processing (e.g., Logie, Gilhooly, & Wynn, 1994; De Rammelaere, Stuyven, & Vandierendonck, 1999). In order to implement this task, we changed our standard food allergy procedure slightly so that the foods constituting a compound meal were presented serially, as the first and second course of the meal, rather than simultaneously. In addition, there was a brief, 2-s interstimulus interval between the first and second courses. In the interference condition we presented briefly three numerals during this interval, one after the other, under the instruction to add up these numbers and retain the total until prompted to supply this total at the end of the trial.

Our assumption in designing this task was that the participants would be still engaged in processing the numbers at the time when the second course was presented and would therefore devote less processing to establishing within-compound associations between the foods of the first and second courses. We also made two further assumptions: first, that mental arithmetic would not interfere with the simple retention of the identity of foods, so that information about the constituents of the meal would be available at the time when the outcome was presented; and, second, that the simple retention of the total number would not interfere with learning about the association between the foods and the outcome. So we assumed that the mental arithmetic task would interfere selectively with the formation of within-compound associations while having little impact on cue-outcome associations. According to the modified theory, therefore, mental arithmetic, by selectively interfering with the formation of within-compounds associations, should reduce retrospective reevaluation.

To test this prediction, we compared causal judgements under the retrospective blocking and generative contingencies using this serial meal procedure in which the treatment foods X and Z were presented as the first course in Stage 1 and target foods B and G as the second course (see Figure 4). As Figure 4 shows, we observed strong retrospective reevaluation in the control, retrospective condition, which was identical to the interference condition except that there was no concurrent mental arithmetic task. Food G from the retrospective generative contingency attracted consistently higher causal ratings than did food B from the retrospective blocking contingency. More importantly, however, the magnitude of retrospective reevaluation was reduced in the interference condition and rendered unreliable. This pattern is exactly the one anticipated by modified associative theory on the assumption that the mental arithmetic task interfered with the formation of within-compound associations between the courses of the meal so that the presentations of the treatment foods X and Z in Stage 2 would, at best, activate only weakly the representation of target foods B and G, respectively.

In fact, we tested this assumption by giving the participants a recognition test after the causal ratings. They were presented with each of the first-course foods and asked to

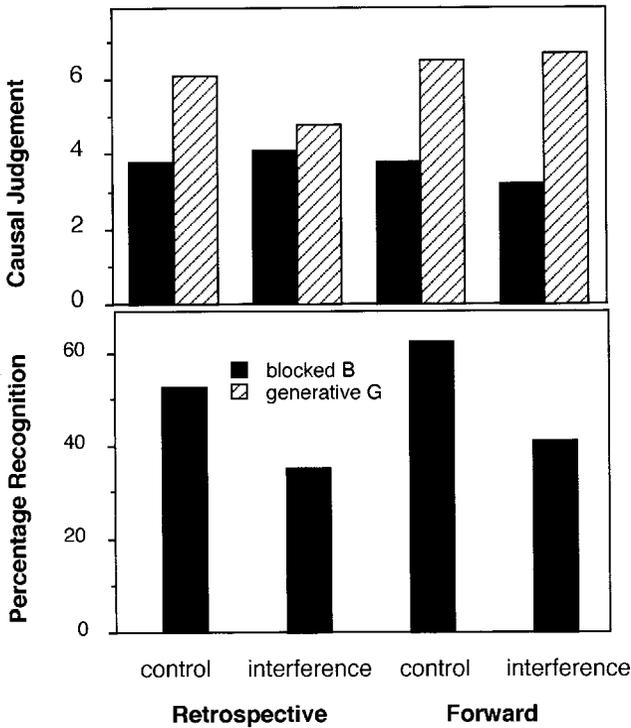
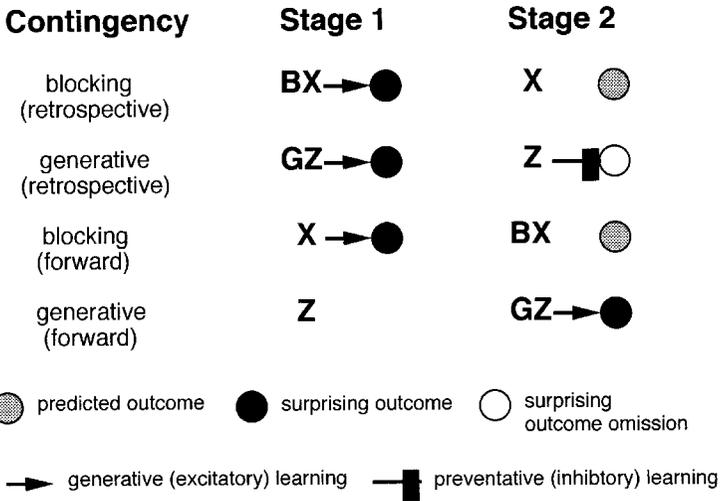


Figure 4. The food-cue–outcome pairings and the standard associative processes for cue presentations in the retrospective and forward blocking and generative contingencies. The top panel displays the mean causal judgements and the bottom panel mean percentage recognition scores for the blocked B and generative G food cues from these contingencies (Aitken et al., 2001).

recognize the second-course food with which it had been compounded from among an array of all the second-course foods. As Figure 4 shows, recognition performance was worse in the interference condition than in the control condition, suggesting that the first-course foods were less likely to retrieve representations of the second-course meal during the Stage 2.

It must be acknowledged, however, that the demonstration that the addition of a second cognitive task produces poorer cue discrimination in a primary task provides, at best, only weak evidence for a specific mechanism of interference. It could well be, for example, that mental arithmetic task interfered with learning the cue–outcome associations. Such a general account of the interference can be evaluated by contrasting the impact of the mental arithmetic task on judgements under the retrospective contingencies with those acquired under forward blocking and generative contingencies.

To recap, these forward contingencies are produced simply by reversing the two stages so that training with the single-course meal precedes training with the two-course meal (see Figure 4). As I have already noted, standard theory predicts that the blocked food B will sustain lower ratings than the generative food G simply because food B is paired with a predicted outcome whereas food G is paired with a surprising outcome. As this basic mechanism of blocking makes no appeal to within-compound associations, the associative analysis predicts that the presence of the interference task should have no impact on the forward blocking of causal learning. By contrast, any general effect of the interference task on sensitivity to the overall contingencies between the food cues and the outcomes should be manifest as strongly in the forward contingencies as in the retrospective ones.

The causal judgements for foods B and G from the forward contingencies favour the modified associative theory in that the inference task had no impact on the magnitude of the blocking effect (see Figure 4). Food G was rated as more causally effective than food B to a comparable extent in the control and interference conditions. Moreover, the fact that the inference task reduced recognition performance to a comparable extent in the retrospective and forward contingencies shows that the insensitivity of blocking to interference was not due to the failure of the mental arithmetic task to impact on learning in the forward contingencies.

In conclusion, the results of this interference study favour not only a role for within-compound associations in retrospective revaluation but also the specific role posited by the modified associative account.

Causal chains

Not all studies of human causal and predictive learning have yielded blocking and retrospective revaluation. For example, Williams, Sagness, and McPhee (1994) found that the form of the cue interaction depends upon whether or not the subjects interpret the cues as exerting independent influences upon the outcome. They observed neither blocking nor retrospective revaluation with neutral instructions in their stock-market task; in order to produce cue interactions, either the instructions setting the scenario or the pretraining had to make explicit that a single cue could exert an independent influence on the outcome. As a result, they suggested that human subjects often treat compound cues as a single configuration (Pearce, 1987) with a causal and predictive status that is relatively independent of the associative strength of the elements of the compound. Such a configural representation is

appropriate when the two putative causes are not individually sufficient but are both necessary to produce the outcome.

The distinction between dependent and independent cues is also important in learning about causal chains. This point can be illustrated by the case of a safety officer who monitors the production processes of a factory via warning lights and gauges, each of which monitors a separate process. The task facing the officer is that of learning which processes cause a safety cut-out in the factory by observing a series of trial runs under the standard two-stage retrospective contingency (see Figure 4). In the first stage, the officer receives a series of compound trials in each of which both a warning light Z and a gauge reading G are paired with the cut-out; then, in the second stage, training is given in which the warning light Z alone occurs without the cut-out. At issue is whether information about the process registered by the warning light should modulate, retrospectively, judgements of the causal status of the processes signalled by the gauge.

As this training instantiates the standard retrospective generative contingency, both a rational analysis and the modified associative theory predict that experience with the warning light Z alone should enhance causal judgements for the process registered by the gauge G . This inference is only warranted, however, if the two processes are independent, putative causes of the cut-out. However, an alternative interpretation of the factory scenario is that the warning light and the gauge register processes that are links in a common causal chain so that one process causes the second to operate, which then, in turn, causes the cut-out. Under this interpretation, it is far less clear how information provided by the warning light should impact upon judgements about the gauge process.

The warning light could be glossed as signalling the terminal link in a $G \rightarrow Z \rightarrow O$ chain (in which O represents the cut-out or outcome). Under this interpretation, the retrospective contingency should reduce the causal status of the gauge G process by demonstrating that the terminal $Z \rightarrow O$ link has been broken. What is less clear, however, is the appropriate inference if the light Z process is seen as the initial link in a $Z \rightarrow G \rightarrow O$ chain. This is especially so if the scenario is one on which information about the state of the gauge G process is no longer available during the second stage. The failure of the light Z process to cause the cut-out could be due to a break either in the first $Z \rightarrow G$ link or in the terminal $G \rightarrow O$ link.

Mark Larkin, Mike Aitken, and I have investigated whether the form of retrospective reevaluation observed using the factory task depends upon the causal structure that is implicated by the scenario. On each trial the subject was presented—on an actual monitor—with two schematic monitors. The left monitor could display one of four warning lights and the right monitor one of four gauges, thereby allowing us to present four different light-gauge compounds. One compound, gauge B and light X , was trained under the retrospective blocking contingency and another pair, gauge G and light Z , under the retrospective generative contingency (see Figure 4). Thus, during the first stage, the BX and GZ compounds were each paired with the safety cut-out on six of eight training trials before, in the second stage, light X alone was consistently paired with the outcome for four trials while the light Z alone was presented without the outcome. A simulated on-off switch on the right-hand monitor, which displayed the gauges, informed the subjects that this monitor was switched off during the second stage, so that the participants had no information about whether the processes registered by the gauges were operative during Stage 2.

The remaining two light-gauge compounds and the light cues from these compounds were presented with the same frequency as the other cues but never paired with the outcome. The function of these cues was to ensure that the subjects attended to the identity of the individual lights and gauges in learning to predict the occurrence of the safety cut-out during training. Although a particular warning light was always compounded with the same gauge, the assignment of a light-gauge compound to each contingency was counterbalanced across subjects.

The instructions given to the independent group ($n = 8$) emphasized that the warning lights and gauges registered independent causal processes, and when finally asked to rate the causal efficacy of the processes on a scale from 100 (safety cut-out will always occur) to 0 (safety cut-out will never occur), the standard retrospective reevaluation effect was observed. The process registered by gauge G from generative retrospective contingency was assigned a higher causal rating than the process registered by gauge B from the retrospective blocking contingency (see Figure 5).

By contrast, exactly the opposite pattern was observed for the second, chain group ($n = 8$). The instructions and training received by this chain group were intended to establish that processes registered by the warning light and the gauge were links in a common chain of events causing the outcome. First, all statements about the independence of the processes were removed from the instructions. Second, the subjects received pretraining designed to

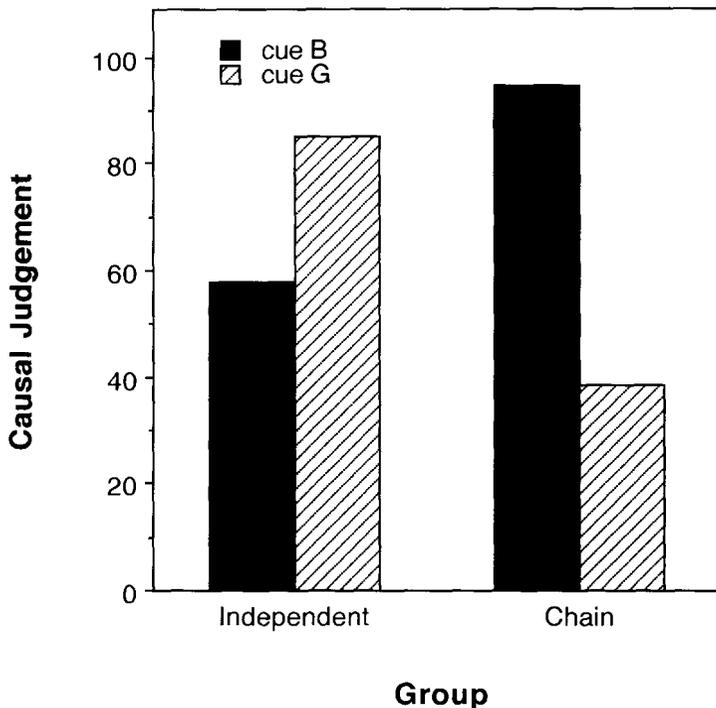


Figure 5. The mean causal judgements for food cues B and G given by the independent and chain groups after training in the factory safety-cut experiment.

establish a causal link between the particular light process and a particular gauge process under a cover-story of testing the warning lights and gauges. The pretraining display on a touch-screen presented the subject with the four warning lights in a row across the top of a touch screen and the four gauges in a diamond pattern below, so that there was no spatial association between any particular light and any particular gauge. At the start of a pretraining trial, the lights and gauges were presented in a greyed-out form. One of the gauges was then illuminated, and the subject's task was to learn by trial and error which warning light to press to cause the gauge to register. In order to get gauge B to register, the subject had to learn to press light X and, similarly, to press light Z to make gauge G register. Four pretraining trials were given with each light-gauge pair in a random order.

Our assumption in designing this pretraining task was that the subjects would learn that the process registered by a warning light activated the process registered by the associated gauge. Finally, to emphasize the cause-effect relationship between the light and gauge processes, on the compound trials of the first stage of revaluation training the onset of the warning light occurred 1,500 ms before the onset of the gauge. Thus, these training conditions were designed to foster a $Z \rightarrow G \rightarrow O$ chain interpretation of the scenario. In all other respects, however, the training received by the chain group was the same as that of the independent group.

Figure 5 shows that these procedural changes produced a complete reversal in the form of the retrospective revaluation.⁶ In the chain group, the causal ratings for gauge G were reduced, rather than enhanced, relative to the ratings for gauge B. In the absence of a control contingency in which the warning light is not trained during the second stage, we cannot determine whether the locus of the revaluation effects lies in the retrospective blocking contingency or the retrospective generative contingency (see Figure 4). The important point, however, is that changing the causal scenario completely reversed the form of the revaluation effect. With the independent scenario, the generative contingency yielded a higher causal rating for the gauge process than the blocking contingency, whereas the chain scenario and pretraining produced exactly the opposite pattern of revaluation.

Clearly, this reversal of judgements is problematic for modified associative theory, which, as it stands, predicts the revaluation pattern observed for the independent group. Before abandoning associative theory, however, it should be noted that the pattern of revaluation demonstrated by the chain group is in fact exactly that anticipated by standard associative theory on the assumption that the same rather than the opposite learning processes are engaged by cue presentations and associatively activated cue representations. A standard associative account of the chain revaluation takes the following form. Both the pretraining with the warning lights and gauges and the compound training establishes an association between light Z and gauge G, so that the presentations of light Z in the second stage activate a representation of gauge G, which is then paired with the surprising omission of the outcome (see Figure 3). Although this is a condition for generative learning according to the modified theory, a simple extension of standard associative theory assumes that pairing an

⁶ A mixed analysis of variance of the ratings for gauges B and G from the retrospective blocking and generative contingencies yields a significant Group X Contingency interaction, $F(1, 14) = 19.78$, $MSE = 704.88$, $p < .01$, and simple main effects analyses establish that the difference in the rating for cues B and G is marginally reliable for the independent group, $F(1, 14) = 4.25$, $p = .06$, and significant for the chain group, $F(1, 14) = 17.88$, $p < .01$.

activated cue representation with outcome omission produces preventative or inhibitory learning, thereby reducing the causal status of gauge G. The claim that learning through cue representations conforms to standard associative processes is not an entirely arbitrary assumption; indeed, standard theory has been routinely applied to cue representations in the explanation of various mediated learning effects in animal conditioning (see Hall, 1996, for a review).

The problem then becomes one of determining, in a principled fashion, when to apply the standard as opposed to the modified theory to retrieved cue representations. Although we cannot be certain about which aspects of the chain training reversed the revaluation pattern, it is notable that this training was designed to establish a causal or predictive contingency between the cues—in fact a contingency that is similar to the one between a cue and an outcome. A fundamental assumption in the associative explanation of causal learning must be that the associations established by such contingencies represent a causal relationship if they are to support judgements of the causal efficacy of a cue. By contrast, the relationship between the cues in the independent contingencies, whether between the courses of a meal in the food-allergy scenario or between warning lights and gauges in the factory scenario, is one of temporal contiguity. Therefore, the within-compound associations established by these contingencies should represent simple temporal contiguity rather than causation.

It is clear, therefore, that if associative theory is to encompass both the interaction between independent causes and the structure of causal chains, two types of associations are required, each with different representational properties: within-compound associations representing simple temporal contiguity and cue (or action)–outcome associations representing a predictive or causal relationship. This distinction is not without precedent in the animal conditioning literature. In their seminal analysis of within-compound learning in conditioning, Rescorla and Durlach (1981) distinguished between what they called within-event learning and between-event learning and suggest that these two forms of associative learning may have different properties. Within the context of the retrospective revaluation of causal judgements, however, the crucial property distinguishing these types of associations is the nature of the learning processes that apply to representations activated by these associations. Whereas the standard learning processes appear to apply to representations activated by causal associations, the inverse processes apply when a representation is activated by a within-compound association.

SUMMARY AND CONCLUSIONS

In this lecture I have presented an analysis of causal learning that reflects my own research interests. As I stated at the outset, my work grew from a concern with the issue of whether instrumental actions are based upon beliefs about the causal effectiveness of these actions in generating desired outcomes. A role for causal cognition in the control of action is supported by the concordance between instrumental performance and the strength of causal beliefs that David Shanks and I (Shanks & Dickinson, 1991) observed across variations in both action–outcome contiguity and contingency.

As the effects of contiguity and contingency in conditioning are explained in terms of associative learning processes, this concordance led, in turn, to an investigation of the role

of these processes in causal learning. Although causal learning exhibits many of the cardinal features of associative processes—blocking, preventative learning, and super-learning—the fact that the causal status of a cue can be changed retrospectively challenges the standard theories. However, associative theory can be modified and extended to encompass retrospective reevaluation by permitting the learning of cue–outcome associations even in the absence of the cue itself. Although a number of modifications have been suggested, they all make two central assumptions: First, that a representation of the absent cue is activated or retrieved on learning episodes when retrospective reevaluation takes place; second, that the form of learning supported by a cue representation retrieved via a within-compound association is the opposite of that accruing to a presented cue. When generative or excitatory learning occurs with a presented cue, the same conditions produce preventative or inhibitory learning with a retrieved cue and, conversely, the conditions that support preventative learning with a presented cue yield generative learning for an activated cue representation. However, our study of the retrospective reevaluation of causal chains suggests that the inversion of the learning processes only applies to cue representations retrieved via within-compound associations representing simply event contiguity. Learning about absent cues whose representations are activated through causal or predictive associations conform to the standard generative and preventative learning processes.

The study of causal reasoning has burgeoned over the last twenty years or so (see Shanks, Holyoak, & Medin, 1996; Sperber, Premack, & Premack, 1995), and I acknowledge that even an elaborated and modified version of associative theory can only claim to be one of many processes supporting the acquisition of causal beliefs and judgements. Associative theory applies only to judgements based upon episodes, whether real or simulated, in which the causal processes are manifest, rather than to causal inferences derived from summary presentations of contingency information or narratives and descriptions (Shanks, 1991). Nor does associative theory address the issue of how domain-specific knowledge, such as that relating to naive physics, (e.g., Baillargeon, Kotovsky, & Needham, 1995; Michotte, 1963; Shultz, 1982), informs causal judgements. But even with these exclusions, a variety of alternative analyses of causal induction are based the experience of event contingencies.

A number of theorists distinguish between the computational and algorithmic levels of analysis (e.g., Cheng, 1997; Perales & Shanks, 2000). Analyses at the computational level seek to specify how asymptotic causal judgements can be computed from contingency information in a way that conforms to the actual judgements given after experience with a contingency. The most sophisticated version of such theories is Cheng's Power PC theory (1997), which uses a basic concept of causal power to constrain judgements based upon contingency information. Causal-model theory (e.g., Waldman, 1996) widens these constraints to include both domain-specific causal knowledge and more abstract, domain-independent models of the nature of different causal processes. Thus, for example, causal-model theory would interpret the different forms of retrospective reevaluation that we observed in the chain and independent groups in terms of the differing causal models that the two groups used to interpret the contingency information.

The main empirical dispute at this level of analysis is whether such computational models yield more accurate descriptions of the judgements observed under different contingencies than do the asymptotic predictions of associative theories (e.g., Cheng, 1997; Lober & Shanks, 2000; Perales & Shanks, 2000; Vallée-Tourangeau, Murphy, Drew, & Baker,

1998)—a dispute that lies outside the scope of this lecture. From a psychological perspective, the main limitation of the computational level of analysis is that it bears only indirectly upon the nature of the processes underlying the acquisition of contingency information, and therefore variables that affect acquisition processes lie outside the scope of these models. Consequently, our finding that treatments that disrupt the formation of within-compound associations impact selectively upon the retrospective reevaluation but not upon forward blocking (Aitken et al., 2001; Dickinson & Burke, 1996) is simply not addressed by computational or causal-model theories.

This finding does, however, discriminate between different algorithmic-level theories of causal and predictive learning. Associative models are obviously prime candidates for such algorithmic processes, and any associative theory that assumes that learning is modulated by the surprisingness of the outcome and allows for the retrieval of cue representations via within-compound associations (e.g., McLaren et al., 1989; Rescorla & Wagner, 1972; Wagner, 1981) can, in principle at least, be modified to explain the forms of causal learning observed in our studies. This is not to say, however, that all associative models, even among those deploying within-compound associations, remain unchallenged by our findings.

For example, Matute, Arcediano, and Miller (1996—see also Shanks & Dickinson, 1988) have suggested that the comparator theory, developed by Miller and his colleague (e.g., Miller & Matzel, 1988) within the context of animal conditioning, might well also govern causal judgements. The basic idea is that all associations, whether they be within-compound or cue–outcome associations, are formed by simple contiguity alone, independently of whether or not the outcome is surprising. So, within this model, all associations represent simple contiguity relationships between events. When a causal or predictive judgement is required for a target cue following compound training, the outcome representation is activated by the target cue via two routes. The first is a direct activation of the outcome representation via the target–outcome association, whereas the second is an indirect activation via the within-compound association with the other cue, the comparator cue, and then by the comparator–outcome association. Comparator theory argues that the judgement about the causal or predictive status of the target cue is determined by the output of a comparison process that reflects the strength of the direct activation of the outcome representation relative to its indirect activation.

Any operation that weakens the capacity of the target to activate the outcome representation *indirectly* should therefore enhance the causal judgement for the target cue. And the indirect activation can be reduced by decreasing the strength of either or both the within-compound association and the comparator–outcome association. These predictions clearly accord with our finding of augmented judgements under the generative retrospective contingency and the dependence of this enhancement on within-compound associations. Comparator theory argues, however, that exactly the same associations should be engaged by judgements based on the forward contingencies. Therefore, the failure of manipulations of the within-compound associations to impact on judgements under forward contingencies is problematic for this type of account.

In conclusion, our analysis suggests that associative processes can support the acquisition of causal beliefs based on experience with event contingencies. This is not to say, however, that associative learning is the sole route to causal knowledge. In certain circumstances, causal judgements can surely be based upon the application of causal models and domain-

specific knowledge to episodically recalled or explicitly represented information about event conjunctions and disjunctions. Indeed, our rational and scientific analysis of causation must be based upon such knowledge. Rather, my claim is that such information is not the only source of contingency information, and, in this respect, it is significant that our procedures typically impose a high mnemonic load by requiring each subject to learn about a number of contingencies concurrently, each of which is often instantiated by multiple cues. Moreover, I should argue that it is surely these conditions that more closely mimic our daily immersion in the causal texture of the world, for which associative learning is ideally adapted. Perhaps the contribution of associative learning to causal knowledge should surprise no one for, over 250 years ago, the greatest of all philosophers and psychologists of causation, David Hume, noted that “’tis sufficient to observe, that there is no relation which produces a stronger connexion in the fancy, and which makes one idea more readily recall another, than the relation of cause and effect betwixt their objects” (Hume, 1739/1888, p. 11).

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.
- Alloy, L.B., & Abramson, L.Y. (1979). Judgment of contingency in depressed and nondepressed students: Sadder but wiser? *Journal of Experimental Psychology: General*, *108*, 441–485.
- Baillargeon, R., Kotovsky, L., & Needham, A. (1995). The acquisition of physical knowledge in infancy. In D. Sperber, D. Premack & A.J. Premack (Eds.), *Causal cognition* (pp. 70–116). Oxford: Clarendon Press.
- Brewer, W.F. (1974). There is no convincing evidence for operant or classical conditioning in adult humans. In W.B. Weimer & D.S. Palermo (Eds.), *Cognition and the symbolic processes* (pp. 1–42). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Chapman, G.B. (1991). Trial order affects cue interaction in contingency judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 837–854.
- Chapman, G.B., & Robbins, S.J. (1990). Cue interactions in human contingency judgment. *Memory and Cognition*, *18*, 537–545.
- Chatlosh, D.L., Neunaber, D.J., & Wasserman, E.A. (1985). Response–outcome contingency: Behavioral and judgmental effects of appetitive and aversive outcomes with college students. *Learning & Motivation*, *16*, 1–34.
- Cheng, P.W. (1997). From covariation to causation: A causal power theory. *Psychological Review*, *104*, 367–405.
- De Rammelaere, S., Stuyven, E., & Vandierendonck, A. (1999). The contribution of working memory resources in the verification of simple arithmetic sums. *Psychological Research*, *62*, 72–77.
- Dickinson, A. (1994). Instrumental conditioning. In N.J. Mackintosh (Ed.), *Animal cognition and learning* (pp. 45–79). London: Academic Press.
- Dickinson, A., & Burke, J. (1996). Within-compound associations mediate the retrospective revaluation of causality judgements. *Quarterly Journal of Experimental Psychology*, *49B*, 60–80.
- Dickinson, A., Shanks, D.R., & Evenden, J.L. (1984). Judgement of act–outcome contingency: The role of selective attribution. *Quarterly Journal of Experimental Psychology*, *36A*, 29–50.
- Dickinson, A., Watt, A., & Griffiths, W.J.H. (1992). Free-operant acquisition with delayed reinforcement. *Quarterly Journal of Experimental Psychology*, *45B*, 241–258.
- Dickinson, A., Watt, A., & Varga, Z.I. (1996). Context conditioning and free-operant acquisition under delayed reinforcement. *Quarterly Journal of Experimental Psychology*, *49B*, 97–110.

- Graham, S. (1999). Retrospective reevaluation and inhibitory associations: Does perceptual learning modulate perception of the contingencies between events? *Quarterly Journal of Experimental Psychology*, *52B*, 159–185.
- Hall, G. (1996). Learning about associatively activated stimulus representations: Implications for acquired equivalence and perceptual learning. *Animal Learning and Behavior*, *24*, 233–255.
- Hammond, L.J. (1980). The effects of contingencies upon appetitive conditioning of free-operant behavior. *Journal of the Experimental Analysis of Behavior*, *34*, 297–304.
- 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.
- Hume, D. (1739/1888). *A treatise on human nature* (L.A. Selby-Brigge, Ed.). Oxford: Clarendon Press.
- Jenkins, H., & Ward, W. (1965). Judgment of contingency between response and outcomes. *Psychological Monographs*, *7*, 1–17.
- Kamin, L.J. (1969). Selective association and conditioning. In N.J. Mackintosh & W.K. Honig (Eds.), *Fundamental issues in associative learning* (pp. 42–64). Halifax: Dalhousie University Press.
- Kaufman, M.A., & Bolles, R.C. (1981). A nonassociative aspect of overshadowing. *Bulletin of the Psychonomic Society*, *18*, 318–320.
- Larkin, M.J., Aitken, M.R.F., & Dickinson, A. (1998). Retrospective reevaluation of causal judgements under positive and negative contingencies. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *24*, 1331–1352.
- Lober, K., & Shanks, D.R. (2000). Is causal induction based upon causal power? Critique of Cheng (1997). *Psychological Review*, *107*, 195–212.
- Logie, R.H., Gilhooly, K.J., & Wynn, V. (1994). Counting on working memory in arithmetic problem solving. *Memory and Cognition*, *22*, 395–410.
- Mackintosh, N.J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review*, *82*, 276–298.
- Matute, H., Arcediano, F., Miller, R.R. (1996). Test question modulates cue competition between causes and between effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 182–196.
- McLaren, I.P.L., Kaye, H., & Mackintosh, N.J. (1989). An associative theory of the representation of stimuli: Applications to perceptual learning and latent inhibition. In R.G.M. Morris (Ed.), *Parallel distributed processing—Implications for psychology and neurobiology* (pp. 102–130). Oxford: Oxford University Press.
- Michotte, A. (1963). *The perception of causality*. London: Methuen.
- Miller, R.R., Barnet, R.C., & Graham, 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., & 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. (1987). A model of stimulus generalization for Pavlovian conditioning. *Psychological Review*, *94*, 61–73.
- 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.
- Perales, J.C., & Shanks, D.R. (2000). Normative and descriptive accounts of the influence of power and contingency on causal judgement. Unpublished manuscript.
- Rescorla, R.A. (1971). Variations in the effectiveness of reinforcement following prior inhibitory conditioning. *Learning and Motivation*, *2*, 113–123.
- Rescorla, R.A. (1991). Associative relations in instructional learning: The Eighteenth Bartlett Memorial Lecture. *Quarterly Journal of Experimental Psychology*, *43B*, 1–23.
- 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 non-reinforcement. In A.H. Black & W.F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). New York: Appleton-Century-Crofts.

- Shanks, D.R. (1985). Forward and backward blocking in human contingency judgements. *Quarterly Journal of Experimental Psychology*, *37B*, 1-21.
- Shanks, D.R. (1987). Acquisition functions in causality judgement. *Learning and Motivation*, *18*, 147-166.
- Shanks, D.R. (1991). On similarities between causal judgements in experienced and described situations. *Psychological Science*, *2*, 341-350.
- Shanks, D.R., & Dickinson, A. (1988). Associative accounts of causality judgment. In G.H. Bower (Ed.), *The psychology of learning and motivation*, Vol. 21 (pp. 229-261). New York: Academic Press.
- Shanks, D.R., & Dickinson, A. (1991). Instrumental judgment and performance under variations in action-outcome contingency and contiguity. *Memory and Cognition*, *19*, 353-360.
- Shanks, D.R., Holyoak, K.J., & Medin, D.L. (1996). *The psychology of learning and motivation: Vol. 34. causal learning*. San Diego, CA: Academic Press.
- Shanks, D.R., Lopez, F.J., Darby, R.J., & Dickinson, A. (1996). Distinguishing associative and probabilistic contrast theories of human contingency judgment. In D.R. Shanks, K.J. Holyoak, & D.L. Medin, D.L. (Eds.), *The psychology of learning and motivation: Vol. 34. Causal learning*. (pp. 265-312). San Diego, CA: Academic Press.
- Shultz, T.R. (1982). Rules of causal attribution. *Monograph of the Society for Research in Child Development*, *47* (1) Serial No. 194.
- Sperber, D., Premack, D., & Premack, A.J. (1995). *Causal cognition*. Oxford: Clarendon Press.
- Sutton, R.S., & Barto, A.G. (1981). Toward a modern theory of adaptive networks: Expectation and prediction. *Psychological Review*, *88*, 135-171.
- Vallée-Tourangeau, F., Murphy, R.A., Drew, S., & Baker, A.G. (1998). Judging the importance of constant and variable candidate causes: A test of Power PC theory. *Quarterly Journal of Experimental Psychology*, *51A*, 65-84.
- Van Hamme, L.J., & Wasserman, E.A. (1994). Cue competition in causality judgements: The role of non-presentation 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 and R.R. Miller (Eds.), *Information processing in animals: memory mechanisms* (pp. 5-47). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Waldman, M.R. (1996). Knowledge-based causal induction. In D.R. Shanks, K.J. Holyoak, & D.L. Medin, D.L. (Eds.), *The psychology of learning and motivation: Vol. 34. Causal learning*. (pp. 47-88). San Diego, CA: Academic Press.
- Wasserman, E.A., & Berglan, L.R. (1998). Backward blocking and recovery from overshadowing in human causal judgment: The role of within-compound associations. *Quarterly Journal of Experimental Psychology*, *51B*, 121-138.
- Wasserman, E.A., Elek, S.M., Chatlosh, D.L., & Baker, A.G. (1993). Rating causal relations: The role of probability in judgment of response-outcome contingency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 174-188.
- Williams, D.A. (1995). Forms of inhibition in animal and human learning. *Journal of Experimental Psychology: Animal Behavior Processes*, *21*, 129-142.
- Williams, D.A., & Docking, G.L. (1995). Associative and normative accounts of negative transfer. *Quarterly Journal of Experimental Psychology*, *48A*, 976-998.
- Williams, D.A., Sagness, K.E., & McPhee, J.E. (1994). Configural and elemental strategies in predictive learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 694-709.