

# A dissociation between causal judgment and outcome recall

CHRIS J. MITCHELL, PETER F. LOVIBOND, and CHEE YORK GAN  
*University of New South Wales, Sydney, New South Wales, Australia*

It has been suggested that causal learning in humans is similar to Pavlovian conditioning in animals. According to this view, judgments of cause reflect the degree to which an association exists between the cause and the effect. Inferential accounts, by contrast, suggest that causal judgments are reasoning based rather than associative in nature. We used a direct measure of associative strength, identification of the outcome with which a cause was paired (cued recall), to see whether associative strength translated directly into causal ratings. Causal compounds AB+ and CD+ were intermixed with A+ and C- training. Cued-recall performance was better for cue B than for cue D; thus, associative strength was inherited by cue B from the strongly associated cue A (augmentation). However, the reverse was observed on the causal judgment measure: Cue B was judged to be less causal than D (cue competition). These results support an inferential over an associative account of causal judgments.

Both human and nonhuman animals need to be sensitive to the causal relationships between events in order to behave adaptively. It is not surprising, therefore, that a common mechanism—associative learning—has been proposed to account for the ability of humans and nonhumans to adapt to changing causal relationships in the environment (Dickinson, Shanks, & Evenden, 1984).

A scenario that has been widely used to investigate the underlying mechanism of causal learning in humans is the *allergist task* (Larkin, Aitken, & Dickinson, 1998). In this task, participants are informed that Mr. X has eaten (for example) cabbage, and they are then told that Mr. X suffered an allergic reaction. On test, participants commonly attribute the allergic reaction to the consumption of cabbage. The most straightforward account of this process in associative terms is that, because the cue (cabbage) was followed by the outcome (the allergic reaction), an association formed between the mental representations corresponding to these two events. Cabbage was judged to cause the allergic reaction because its presentation produced activation of the allergic reaction representation.

A range of findings support this associative account of causal learning (see De Houwer & Beckers, 2002, for a review), and the most important of these is cue competition (see, e.g., Chapman & Robbins, 1990). In cue competition, Mr. X might consume two different meals, cabbage and bread (compound cue AB) and pizza and beans (compound cue CD), and suffer an allergic reaction (+) following both meals (AB+ and CD+). If food cue A is also observed to be allergenic when presented alone (A+), but food C is safe (C-), then the causal rating of D

will be higher than that of B on test. In associative learning terminology, the outcome on AB+ and CD+ trials can support only a limited amount of associative strength, for which each cue must compete. Therefore, A+ trials reduce, and C- trials increase, the amount of strength that remains to be made available to the partner cues B and D, respectively.

In general, the associative account makes the strong prediction that patterns of causal ratings will reflect the degree to which the allergic reaction outcome node is activated when a food cue is presented. Thus, outcome activation, or associative strength, could be measured independently to test this prediction. One method that has been used to assess associative strength in past studies is cued recall. Aitken, Larkin, and Dickinson (2001; see also Melchers, Lachnit, & Shanks, 2004) used cued recall to index the strength of an association between two food cues in their causal judgment experiment. Thus, a target food was presented on test, and participants were required to identify from a separate list the food that had appeared with the target during training. Accuracy on this task was taken to indicate the strength of the association between the two foods. Mitchell, Lovibond, Minard, and Lavis (in press) used a similar cued-recall task to measure the strength of association between a food cue and an outcome in the allergist task. In this design, a single cue was paired with an allergic reaction when presented alone (A+) early in training. Later, a compound of foods A and B was also shown to be causal (AB+). A second causal compound (CD+) served as a control. In order to allow a cued-recall test of associative strength in this study, foods were paired with a range of different allergic reaction outcomes. On test, participants were required to identify the particular outcome with which the food had been paired and give a causal rating. Performance on the cued-recall task was poorer for cue B than for D, and causal ratings of B were lower than those of D. Thus, cue A “blocked”

---

This work was supported by a grant from the Australian Research Council. Correspondence relating to this article should be addressed to C. J. Mitchell, School of Psychology, University of New South Wales, Sydney, NSW 2052, Australia (e-mail: chris.mitchell@unsw.edu.au).

learning about cue B on AB+ trials. Blocking is commonly observed in animals and is considered a hallmark of associative learning.

Evidence for a failure to encode the B–outcome relationship in the A+/AB+ design above would appear to be strong evidence in favor of the associative account of causal judgments, and of cue competition in particular. However, there is also evidence that cue competition can arise through inferential reasoning processes (see, e.g., De Houwer, Beckers, & Glautier, 2002; Lovibond, Been, Mitchell, Bouton, & Frohardt, 2003). For example, in Lovibond et al.'s blocking study, half of the participants learned early in training that a combination of two causal cues (e.g., X+ and Y+) produced a stronger reaction when combined (XY++). This was termed an *additive* cue combination rule. The remaining participants were given the *nonadditive* cue combination rule, in which two causal cues (X+ and Y+) were presented together and the standard reaction was observed (XY+). Stronger blocking was observed in the additive case; only the participants in that group could reason that B was not causal because, if B were causal (B+), then the AB compound would have produced a strong reaction (AB++).

The present experiment focused on the idea that cue competition might arise through inferential reasoning processes even in the absence of any cue combination rules. In our AB+/CD+/A+/C– design, to reason that D, but not B, is causal (cue competition), participants only need to make the intuitive assumption that at least one cue in a causal compound (AB+ and CD+) is causal. Thus, if A is found to be causal (A+), then cue B need not be (B–). In addition, if C is shown to be noncausal (C–), then D certainly ought to be a cause (D+). That is, the higher causal rating of D than of B on test may be the result of propositional reasoning and not of a low-level associative mechanism.

What does the inferential account predict with respect to the relationship between cued recall and causal judgment? Clearly, to make an inference, it is necessary to remember the information upon which that inference is to be made. Thus, the inferential account would be supported over the associative account if cue competition were observed in the AB+/CD+/A+/C– design, despite the absence of any difference between cues B and D on a cued-recall task. Such a dissociation, however, would be quite weak evidence for the inferential account, since the cued-recall task might merely be a less sensitive test of associative strength than the causal judgment task. The inferential account would be more strongly supported if cue D were rated more causal than cue B, despite poorer cued-recall performance to D than to B. The present experiment sought evidence for just such a dissociation; we attempted to reverse the standard cue competition effect, as measured in the cued-recall task, and we then observed the impact of this reversal on the causal attributions made for those cues.

Past animal studies have suggested that, although blocking is often observed in taste aversion learning in rats (see, e.g., Gillan & Domjan, 1977), A+ trials can increase, rather than reduce, the associative strength accrued by cue B on AB+ trials (e.g., Batsell, Paschall, Gleason, & Batson, 2001). Batsell et al. concluded that this *augmentation* occurred when a strong within-compound association existed between cues A and B. Thus, the high associative strength of A was able to spread to its partner cue B via the within-compound association. Following this suggestion, the compounds we used in the present study were chosen to have elements that were already highly associated at the start of training (e.g., fish and chips). In addition, we gave very few compound training trials (AB+ and CD+) but many element trials (A+ and C–). We expected these two parametric manipulations to result in responding to cues B and D that was largely determined by inheritance of associative strength from the elements with which they were paired (A and C), rather than by the competition for associative strength that might be expected to occur on (the very rare) compound training trials.

According to associative theory, if our measure of associative strength (cued recall) reveals augmentation rather than cue competition, augmentation should be observed in the causal judgment task. The inferential view allows that cue competition might be observed on the causal judgment measure, despite an augmentation effect in cued recall.

## METHOD

### Participants

The participants were 21 undergraduate psychology students from the University of New South Wales who volunteered to take part in this experiment in return for course credit.

### Apparatus and Stimuli

Instructions and daily allergy trials were presented on a computer screen. The participants progressed through the screen pages and responded to the test questions by using the mouse to click on buttons presented on the screen. The cues were eight well-known food pairs: *macaroni and cheese*, *peaches and cream*, *bacon and eggs*, *bread and butter*, *tea and coffee*, *salt and pepper*, *fish and chips*, and *spaghetti and meatballs*. These foods were randomly assigned to the roles of eight compounds, A1B1–A4B4 and C1D1–C4D4, in the experimental design. The outcomes were four fictitious allergic reactions: *daryosis*, *chloristine*, *plonthema*, and *xianethis* (Outcomes 1–4, respectively).

Each trial screen depicted the words *At this meal, Mr X eats:* and, beneath them, the names of one or two foods. The participants were required to predict the outcome of each trial by clicking on one of five buttons presented at the bottom of the screen. Four of the buttons corresponded to the fictitious allergic reactions, and the fifth allowed the participants to respond *no allergy*. The correct outcome was then presented on a feedback screen. Allergic reactions were presented in a red font, *no allergy* in a green font.

For the test, all foods were presented on the same screen. Beneath each food appeared the four allergic reactions. The *no allergy* option was not available on test. The participants were required to click on one of the four reactions to indicate which they thought Mr. X suffered from after consuming that food during training. For example,

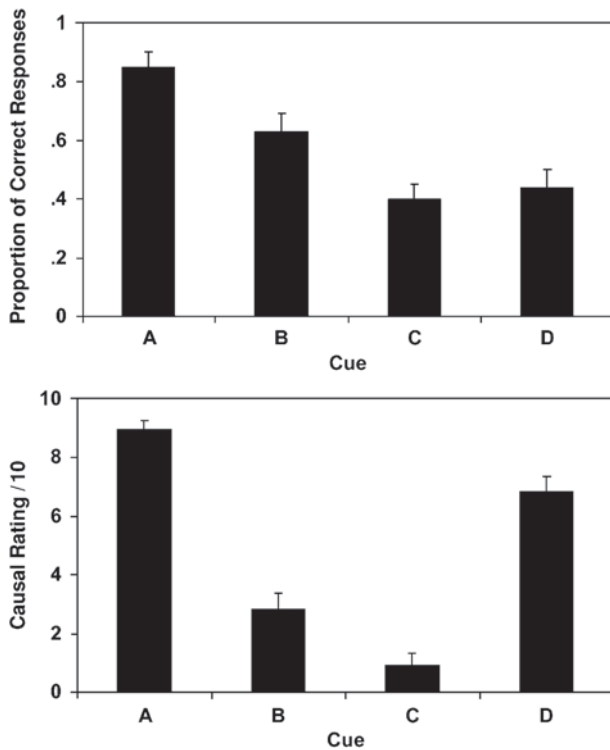


Figure 1. The top panel shows the proportion of correct responses made to each cue type, A–D, and the bottom panel shows the causal judgments for the same cues, averaged across participants. The error bars indicate standard errors of the means.

the outcome *xianethis* might be indicated in response to the cue *fish*. Once the participant clicked on *xianethis*, a Likert scale appeared on the right-hand side of the screen with the question *To what extent did fish cause xianethis?* The scale ranged from 0 (*not at all*) on the left to 6 (*completely*) on the right, with seven evenly spaced anchor points. Responses could be made anywhere on the scale.

### Design

During training, the four compounds A1B1–A4B4 were followed by Outcomes 1–4, respectively, as were compounds C1D1–C4D4. These 8 compound training trial types were presented twice each, producing 16 compound training trials in total. There were also 8 element trial types, presented six times each. The elements A1–A4 were paired with Outcomes 1–4, respectively, and the elements C1–C4 were each followed by no reaction. Thus, there were 48 element training trials in total. The 64 total trials (both compound training and element training) were presented intermixed in a random order, with no filler trials. All 16 foods, A1–D4, were then presented on the test screen.

### Procedure

All instructions were presented on the computer screen. The participants were first informed about the general nature of the allergic task. In particular, they were asked to attend to, and remember, both which outcome occurred after each food and to what extent the food caused that outcome. Two examples of feedback were given in the instruction phase: *neuritis* (an outcome not used during training) and *no allergy*.

The 64 training trials were presented with no break and were immediately followed by the test screen. The participants were allowed to make outcome predictions and to move from the feedback screen of one trial to the start of the next trial at their own pace. On the

single test screen, the participants were required to recall the allergic reactions that followed each food and to make a causal rating of each food. They were permitted to change their choices as many times as they wished. Pressing the *continue* button on the test screen ended the test and the experiment.

## RESULTS

Element training proceeded as expected. Across all participants, an allergic reaction of some kind was predicted to occur on 96.4% of the sixth and final A+ trials. In addition, the exact type of allergy that would occur was correctly identified on 95.1% of these trials. In contrast, an allergic reaction was predicted to occur on only 6.0% of the last C– trials. Thus, participants learned to discriminate A+ from C–. The pattern was more complicated with respect to the compound trials. An allergic reaction was predicted to occur on 97.6% of the second (and final) AB+ trials. The correct type of allergic reaction was predicted on 85.4% of those trials. Predictions on CD+ trials were less accurate. Participants correctly predicted that an outcome would occur on 85.7% of these trials, but the correct outcome was predicted on only 40.3%. A contrast comparing performance on the two types of compound trial revealed that the correct outcome was identified less often on the last CD+ trial than on the last AB+ trial [ $F(1,20) = 48.58$ ,  $MS_e = 0.051$ ]. It appears that predictive accuracy on AB+ trials benefited from A+ training trials.

The responses to cues A1–A4 were combined for analysis of the test data and will be referred to as responses to cue A. The same is also true of cues B1–B4, C1–C4, and D1–D4. Accuracy on the cued-recall task is shown in the top panel of Figure 1. Recall of the outcome paired with cue A was better than that of the outcome paired with cue C [ $F(1,20) = 36.41$ ,  $MS_e = 0.056$ ]. In addition, performance to cue B was better than that to cue D [ $F(1,20) = 10.80$ ,  $MS_e = 0.05$ ]. Thus, the improvement on the cued-recall task resulting from the added A+ trials transferred to cue B. The causal ratings of cues A–D, including those whose outcomes were not correctly identified, are shown in the lower panel of Figure 1. Causal ratings of the cues trained separately correspond closely to the training data: Cue A was attributed higher causal strength than was cue C [ $F(1,20) = 224.28$ ,  $MS_e = 7,619.58$ ]. However, cue B was attributed lower causal strength than was cue D [ $F(1,20) = 18.81$ ,  $MS_e = 5,188.63$ ]. These data reveal a cue competition effect: Increasing the causal value of A (relative to C) decreased that of its partner B (relative to D).

## DISCUSSION

Participants given AB+/CD+/A+/C– training were better able on test to recall the outcome that followed cue B during training than they were the outcome that followed cue D. The reverse pattern was observed on the causal judgment measure: Cue B was rated as less causal than cue D on test. Associative theory predicts that the attribution of causality to a cue is a consequence of the

ability of that cue to excite activation of the outcome representation. To the extent that the cued-recall task measures the level of outcome activation when a cue is presented, these data challenge the associative account of causal judgments.

The effect in cued recall is similar to the augmentation effect observed by Batsell et al. (2001) and is the reverse of our previous findings. Mitchell et al. (in press) presented A+ trials prior to AB+ and CD+ compound trials and observed poorer cued-recall performance to B than to D (cue competition). We think it likely that the strong within-compound associations in the present task were responsible for this difference. The animal conditioning literature suggests that strong within-compound associations reduce, and in extreme cases can reverse, cue competition (Batsell et al., 2001; see Durlach, 1989, for a review of some other relevant studies). Furthermore, the data are consistent with the idea, common in the memory literature, of spreading semantic activation (see, e.g., Collins & Loftus, 1975). According to these models, presentation of the B cue on test will activate both the outcome and the strongly associated cue A. Activation will then spread from cue A to the outcome, further enhancing outcome recall. Presentation of cue D will activate the outcome less strongly, because its associate C is less able to activate the outcome than is cue A. Of course, this process will attenuate or even reverse any cue competition that might otherwise have occurred.

The role suggested for within-compound associations in causal judgments is quite different; they are thought to enhance, rather than reverse, cue competition. In particular, when the element trials (A+ and C-) follow the compound trials (AB+ and CD+), strong within-compound associations are necessary to observe cue competition (Aitken et al., 2001; Melchers et al., 2004). According to some associative accounts, the explicit absence of the representation of B will somehow be activated on A+ trials and be followed by the outcome. The pairing of the absent B and the outcome is then thought to reduce, rather than increase, the associative strength of B. This idea has been suggested in a variety of forms (Dickinson & Burke, 1996; Markman, 1989; Van Hamme & Wasserman, 1994).

Thus, an associative account exists of the role of within-compound associations in augmentation and cue competition following AB+/CD+/A+/C- training. However, the associative view has difficulty explaining why the same cues produce opposite outcomes on two measures of associative strength, cued recall and causal judgment, within the same experiment. Clearly, the two outcomes reflect different processes. Matute and her colleagues (see, e.g., Matute, Vegas, & De Marez, 2002) have argued that once participants have acquired information in training, they are able to use it quite flexibly, depending on the test questions. The present data support this view. In Matute's experiments, it was assumed that participants had encoded all associative relationships. The present experiments show that associative strength itself can be dissociated from judgments based on that information.

To account for the present dissociation, we propose an inferential reasoning process in which associative learning (or memory for cue–outcome relationships) is necessary but not sufficient to produce causal judgments. For example, three relationships are relevant to the status of cue B: (1) cue B was presented in a compound with A, (2) this compound was followed by the outcome, and (3) A was paired with the outcome when presented alone. Encoding of these three relationships will allow cue B to activate the outcome in two ways, both directly and via its association with cue A, so cued-recall performance will be high. However, the same three relationships will also allow participants to draw the inference that it was cue A, not B, that caused the outcome on AB+ trials. If these three associations had not been encoded, such an inference would not have been drawn.

In fact, this view has something in common with the “comparator” model proposed by Miller and colleagues (Miller & Matzel, 1988). The *comparator* is a cue that appeared with the target in training and is associated with it. According to this associative model, responding to cue B is lower than responding to cue D because the associative strength of its comparator, A, is higher than the associative strength of D's comparator. The dissociation in the present data can, in principle, be explained in the following way. The cued-recall measure might reflect the raw associative strength of the target cues, with Bs higher than Ds. The causal judgment might then reflect the output of the comparator: The associative strength of B (although high) is lower than that of A, and the strength of D (although low) is higher than that of C. However, there is no reason to suppose that cued recall should reflect raw associative strength rather than the output of the comparator. In fact, there is no mechanism within the comparator model by which raw associative strength can be expressed in behavior at all. The comparator model would, therefore, require further development to account for the present data.

Our inferential account presupposes that cued recall reflects associative activation and that causal judgment reflects inferential reasoning. We could also imagine that the reverse is true, however, so that the cued-recall task requires an inference and causal ratings are based on associative activation of the outcome node. That is, when B is presented in the cued-recall task, participants might infer that, because B was presented with A, and A was followed by *xianethis*, B must also have been followed by *xianethis*. For cue B to then be given a low rating on the causal judgment measure, the *xianethis* outcome must be identified in the absence of any activation of the outcome node. This paradoxical outcome would appear to require two separate systems that operate independently. It may be the case that an episodic memory system allows correct outcome identification in cued recall and an associative system determines causal judgments, which would be a testable (if unparsimonious) hypothesis. If we assume that inferences take time, but that associative activation is relatively automatic, in that case time pressure at test would be expected to disrupt cued-recall performance



but not causal judgments. Our inferential account, on the other hand, predicts that causal judgments will be more strongly affected by time pressure. This is a line of research we are presently pursuing.

In summary, we observed a dissociation between cued recall and causal judgment that is difficult to explain within an associative framework. We propose that associative or relational information is necessary but not sufficient to attribute cause. Thus, in very simple training conditions, such as A+ training, causal ratings will directly reflect the degree to which the cue and outcome are associated in memory. However, in more complex designs such as the one used here, a strong cue–outcome association can exist in the absence of any causal attribution.

## REFERENCES

- 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.
- BATSELL, W. R., PASCHALL, G. Y., GLEASON, D. I., & BATSON, J. D. (2001). Taste preconditioning augments odor-aversion learning. *Journal of Experimental Psychology: Animal Behavior Processes*, **27**, 30-47.
- CHAPMAN, G. B., & ROBBINS, S. J. (1990). Cue interaction in human contingency judgment. *Memory & Cognition*, **18**, 537-545.
- COLLINS, A. M., & LOFTUS, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, **82**, 407-428.
- DE HOUWER, J., & BECKERS, T. (2002). A review of recent developments in research and theories on human contingency learning. *Quarterly Journal of Experimental Psychology*, **55B**, 289-310.
- DE HOUWER, J., BECKERS, T., & GLAUTIER, S. (2002). Outcome and cue properties modulate blocking. *Quarterly Journal of Experimental Psychology*, **55A**, 965-985.
- 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., & EVENDEN, J. (1984). Judgement of act–outcome contingency: The role of selective attribution. *Quarterly Journal of Experimental Psychology*, **36A**, 29-50.
- DURLACH, P. J. (1989). Learning and performance in Pavlovian conditioning: Are failures of contiguity failures of learning or performance? In S. B. Klein & R. R. Mowrer (Eds.), *Contemporary learning theories: Pavlovian conditioning and the status of traditional learning theory* (pp. 19-59). Hillsdale, NJ: Erlbaum.
- GILLAN, D. J., & DOMJAN, M. (1977). Taste-aversion conditioning with expected versus unexpected drug treatment. *Journal of Experimental Psychology: Animal Behavior Processes*, **3**, 297-309.
- 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, & Cognition*, **24**, 1331-1352.
- LOVIBOND, P. F., BEEN, S.-L., MITCHELL, C. J., BOUTON, M. E., & FROHARDT, R. (2003). Forward and backward blocking of causal judgment is enhanced by additivity of effect magnitude. *Memory & Cognition*, **31**, 133-142.
- MARKMAN, A. B. (1989). LMS rules and the inverse base-rate effect: Comment on Gluck and Bower (1988). *Journal of Experimental Psychology: General*, **118**, 417-421.
- MATUTE, H., VEGAS, S., & DE MAREZ, P.-J. (2002). Flexible use of recent information in causal and predictive judgments. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **28**, 714-725.
- MELCHERS, K. G., LACHNIT, H., & SHANKS, D. R. (2004). Within-compound associations in retrospective revaluation and in direct learning: A challenge for comparator theory. *Quarterly Journal of Experimental Psychology*, **57B**, 25-53.
- 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). Orlando, FL: Academic Press.
- MITCHELL, C. J., LOVIBOND, P. F., MINARD, E., & LAVIS, Y. (in press). Forward, but not backward, blocking in human learning sometimes reflects the failure to encode a cue–outcome relationship. *Quarterly Journal of Experimental Psychology*.
- VAN HAMME, L. J., & WASSERMAN, E. A. (1994). Cue competition in causality judgments: The role of nonpresentation of compound stimulus elements. *Learning & Motivation*, **25**, 127-151.

(Manuscript received September 13, 2004;  
revision accepted for publication February 22, 2005.)