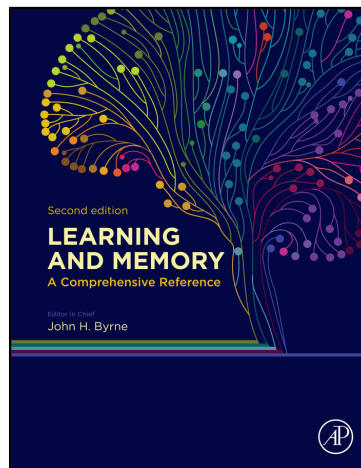


**Provided for non-commercial research and educational use.  
Not for reproduction, distribution or commercial use.**

This article was originally published in *Learning and Memory: A Comprehensive Reference, 2nd edition*, published by Elsevier, and the attached copy is provided by Elsevier for the author's benefit and for the benefit of the author's institution, for non-commercial research and educational use including without limitation use in instruction at your institution, sending it to specific colleagues who you know, and providing a copy to your institution's administrator.



All other uses, reproduction and distribution, including without limitation commercial reprints, selling or licensing copies or access, or posting on open internet sites, your personal or institution's website or repository, are prohibited. For exceptions, permission may be sought for such use through Elsevier's permissions site at:

<https://www.elsevier.com/authors/author-schemas/artwork-and-media-instructions>

From Blaisdell, A.P. (2017) Cognitive Dimension of Operant Learning. In: Menzel, R. (ed.), *Learning Theory and Behavior*, Vol. 1 of *Learning and Memory: A Comprehensive Reference, 2nd edition*, Byrne, J.H. (ed.). pp. 85–110. Oxford: Academic Press. <http://dx.doi.org/10.1016/B978-0-12-809324-5.21007-9>  
ISBN: 9780128051597

Copyright © 2017 Elsevier Ltd. All rights reserved.  
Academic Press

## 1.06 Cognitive Dimension of Operant Learning<sup>☆</sup>

Aaron P Blaisdell, University of California – Los Angeles, Los Angeles, CA, United States

© 2017 Elsevier Ltd. All rights reserved.

<b>1.06.1</b>	<b>Thorndike and the Law of Effect</b>	85
<b>1.06.2</b>	<b>Tolman's Purposive Psychology</b>	87
<b>1.06.3</b>	<b>Operant Behavior: Goal Directed Versus Habitual</b>	87
1.06.3.1	Behavioral Dissociations	88
1.06.3.1.1	Outcome Devaluation	88
1.06.3.1.2	Manipulations of the Response–Outcome Contingency	89
1.06.3.1.3	Outcome Expectation Drives Response Variation	90
1.06.3.2	Neurobiological Dissociations	91
<b>1.06.4</b>	<b>Stimulus–Outcome Associations in Instrumental Discrimination Learning</b>	92
<b>1.06.5</b>	<b>Stimulus–Stimulus Associations in Instrumental Discrimination Learning</b>	94
<b>1.06.6</b>	<b>Agency</b>	96
1.06.6.1	Intentional Psychology: Beliefs and Desires	96
1.06.6.2	Animals as Free Agents	96
<b>1.06.7</b>	<b>Interventions and Causal Reasoning</b>	97
1.06.7.1	Making Things Happen	97
1.06.7.2	Seeing Versus Doing	97
1.06.7.2.1	Seeing Versus Doing in Children	99
1.06.7.2.2	Seeing Versus Doing in Rats	100
<b>1.06.8</b>	<b>Tool Use: From Crow to Cro-Magnon</b>	102
1.06.8.1	Tool Use in Primates	102
1.06.8.2	Tool Use in Corvids	104
1.06.8.3	Tool Use by Humankind	107
<b>Acknowledgments</b>		108
<b>References</b>		108

### 1.06.1 Thorndike and the Law of Effect

A rat presses a lever and quickly scurries to the food cup to collect its meal. A crow picks up a piece of newspaper – uncovering a morsel of food dropped by a passerby. A child lifts the lid off of a cookie jar and takes a cookie. These are the types of acquired behaviors that a theory of instrumental learning seeks to explain. Instrumental learning is the acquisition of a new response through reinforcing feedback. Thorndike (1898, 1911) provided the first general theory of instrumental learning with his Law of Effect. The Law of Effect is a simple trial-and-error learning model in which the actions an individual makes in the presence of a particular stimulus or context are strengthened or weakened depending on the consequences of those actions. Actions followed by desirable consequences—such as attainment of food rewards or escape from aversive situations—become strengthened in that stimulus context, while actions followed by undesirable consequences—such as a loss of food or attainment of an aversive stimulus—become weakened in that context. His theory grew out of experiments on escape learning in animals – most famously cats. In these experiments he studied the acquisition of a new behavior in a controlled manner. A cat was placed in a cage which Thorndike called a puzzle box (Fig. 1) with food located in view just outside. The cat could escape the box by manipulating a device.

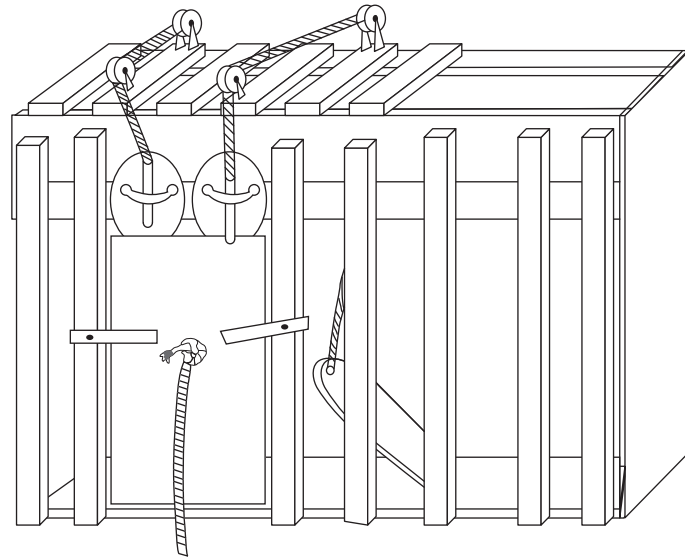
Various puzzle boxes were rigged so that each one required a different manner of escape, such as pushing the door aside, pressing a lever, pulling a string, or a series of these behaviors. Thorndike repeatedly placed the cat inside a puzzle box and observed the latency to escape. As the association between the stimulus (puzzle box) and the effective escape response increased, the latency for the cat to escape diminished (Fig. 2).

What struck Thorndike was that the cat initially tried all manners of escape – trying to squeeze through the bars, pawing at the door, clawing and biting at things in the box. The cat would “...strive instinctively to escape from confinement. The vigor with which it struggles is extraordinary. For eight or ten minutes it will claw and bite and squeeze incessantly” (Thorndike, 1898, p. 13).

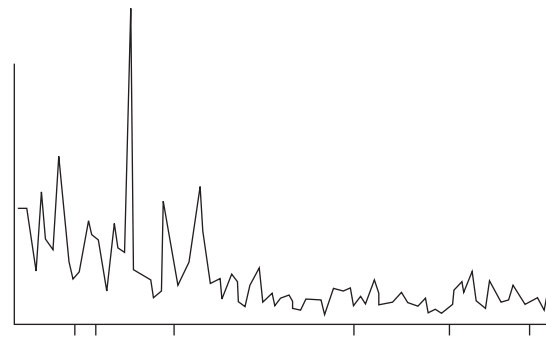
These strivings, however, diminished upon repeated exposures to the box as the cat learned which actions were strengthened and which were weakened. The cat became more successful over repeated trials – successful escape responses that at first had been

<sup>☆</sup>*Change History:* March 2016. AP Blaisdell updated the text and references to this entire chapter, updated keywords and abstract, added new figures 9, 11, and 12, and added two new sections: “Stimulus–Outcome associations in instrumental discrimination learning” and “Stimulus–Stimulus associations in instrumental discrimination learning”.

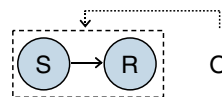
This is an update of A.P. Blaisdell, 1.10 – Cognitive Dimension of Operant Learning. In *Learning and Memory: A Comprehensive Reference*, edited by John H. Byrne, Academic Press, Oxford, 2008, Pages 173–195, ISBN 9780123705099, <http://dx.doi.org/10.1016/B978-012370509-9.00184-4>.



**Figure 1** One of the puzzle boxes used by Thorndike to study the acquisition of new behaviors in cats. From Thorndike, E.L., 1911. *Animal Intelligence: Experimental Studies*. Macmillan, New York.



**Figure 2** Escape times as a function of trials for one cat in a puzzle box experiment. From Thorndike, E.L., 1911. *Animal Intelligence: Experimental Studies*. Macmillan, New York.



**Figure 3** Theoretical associative structure underlying instrumental conditioning. *O*, outcome; *R*, response; *S*, stimulus. *Solid arrow* indicates hypothetical unidirectional excitatory association between the *S* and *R*. The *dashed-line box* envelopes the content of learning. The *dotted arrow* indicates that *O* only plays a role in reinforcing the strength of the *S*–*R* association, but does not enter into an association with either *S* or *R*.

unleashed fortuitously in the random fury of its strivings became more systematically employed, while those behaviors that led to no release died away. It was this transformation from seemingly random behaviors to successfully organized ones that led Thorndike to formulate the Law of Effect.

Importantly, observations of the behavior of the cat during and after learning an escape response led Thorndike to conclude that the cat did not understand the relationship between its behavior and the consequence. Rather, the cat appeared to blindly engage in the trained action whenever it was placed in the box. Thus, in Thorndike's framework the consequences of action play an important role in the strengthening or weakening of behavior, but those consequences do not themselves enter into an association with the action or with the prevailing stimulus conditions. Rather, the subject only learns an association between the stimulus context (*S*) and the response (*R*) (Fig. 3).

Many other behaviorists, perhaps Watson and Skinner being among the most famous, had their own take on S–R psychology. The behaviorist ideology eschewed references to mental terms, such as expectations, wants, and desires, and attempted to describe all acquired behavior through the objective lens of stimulus input and response output (Epstein, 2014; Skinner, 1966), with outcomes and consequences of behavior serving only to strengthen or weaken S–R relationships. Thus, to

characterize the learning by a cat in a puzzle box or by a rat in a maze, the cat (or rat) is said to engage in a series of conditioned responses (CRs) (muscle twitches) when they are in a certain context. The fact that these muscle twitches typically bring about a consequence is irrelevant from the cat's (or rat's) point of view. That is, the behavior of the cat or the rat is not goal directed. If one were to ask the behaviorist "Why did the chicken cross the road?" the reply would be "Because the chicken had crossed the road a number of times in the past, and each time the crossing was followed by a satisfying outcome; whereas each failure to cross the road was followed by an unsatisfying outcome. Thus, when in the presence of the road, the experienced chicken will cross to the other side." Likewise, a child learns to lift the lid off of a cookie jar because doing so in the past has repeatedly been followed by a satisfying outcome. If the child were to be asked why they are lifting the lid off of the cookie jar, the behaviorist would expect a reply of "I dunno."

### 1.06.2 Tolman's Purposive Psychology

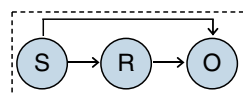
But does it not strike us as being exceedingly odd to interpret the child's act of lifting the lid off the cookie jar as *not* being due to the child's desire for a cookie – as the behaviorist would have it? Tolman agreed, and suggested, that actions are performed for some purpose, that is, they are driven by expectations of an outcome (Tolman, 1932). Experiments conducted in Tolman's lab led to this conclusion. In one example, rats were given nine trials on which they learned to run down a maze and collect a desirable food reward—bran mash—in the goal box. On the 10th trial, the rats found a less desirable sunflower seed in the goal box rather than the bran mash. On the 11th trial, the rats ran down the maze more slowly than they had on the preceding 10 trials. More importantly, they even ran slower than a group of rats that had always received sunflower seeds at the end of the maze. In contrast with behaviorist theory, and with simulations of the Law of Effect (Blaisdell et al., 2016), the rats appeared to have learned during the first nine trials to expect bran mash when they reached the goal box. The violation of this expectation on the 10th trial caused a dramatic and immediate shift in their behavior (Flaherty and Grigson, 1988). Tolman used the behaviorist's controlled methodology and strict adherence to empirical validation to support cognitive processes as variables that intervene between stimulus and response. As a result, cognitive theories of animal learning and behavior have again become fashionable (see Blaisdell, 2003, 2009; Bolles, 1972; Konorski and Miller, 1936; Mackintosh and Dickinson, 1979; Tolman, 1932, for theories of R–O associations in instrumental learning).

### 1.06.3 Operant Behavior: Goal Directed Versus Habitual

In the past few decades, there has been a resurgence in the analysis of goal-directed behavior in psychology according to which an individual's actions are motivated by the outcome representation with which they are associated. Beyond being merely a catalyst for learning, the consequence of a response can become part of what is learned. The response–outcome (R–O) association allows the outcome to motivate the response. That is, the subject *expects* to receive O if it engages in R. A more contemporary view holds that all three elements of an instrumental learning event, S, R, and O, can be bound by associations (Fig. 4).

The claim that instrumental behavior can be goal directed should not be taken to deny the existence of non-goal-directed, habitual behavior. Much of our behavior is clearly produced without the aid of any explicit representation, such as when we walk from our car to our doorstep. We may carry out such a task almost perfectly in complete darkness, or if we are otherwise distracted—such as when we are engaged in conversation with a visitor. Likewise, it is said that one way to catch a sleep walker in the act is by rearranging the furniture in their dwelling. Presumably they will wake up when they bump into furniture that is now blocking a previously unobstructed pathway. This works in rats as well. After being well trained to navigate a maze, rats would bump into the end wall when the length of the maze alley was unexpectedly shortened (Carr and Watson, 1908). Even when we are aware that we have rearranged the furniture in a room, it usually takes time to stop going to the corner where the easy chair used to be to have a seat. Discovering the empty corner quickly reminds us of our error. Old habits die hard, as the saying goes.

To be theoretically useful, there must be a way to empirically distinguish habitual (S–R) from goal-directed (R–O) behavior. According to several theorists (Colwill and Rescorla, 1986; Dickinson and Balleine, 1994), behavior is said to be goal directed if it is mediated by (1) the instrumental contingency between the action and the outcome, and (2) the value of the outcome. This operational definition is consistent with the two lines of evidence in support of R–O associations: posttraining changes in the value of the outcome and manipulations of the R–O contingency.



**Figure 4** Theoretical associative structure underlying instrumental conditioning. O, outcome representation; R, response representation; S, stimulus representation. Arrows indicate hypothetical unidirectional excitatory associations between representations. The dashed-line box envelopes the contents of learning. Associations are hypothesized to form between S and R, between S and O, and between R and O.

### 1.06.3.1 Behavioral Dissociations

#### 1.06.3.1.1 Outcome Devaluation

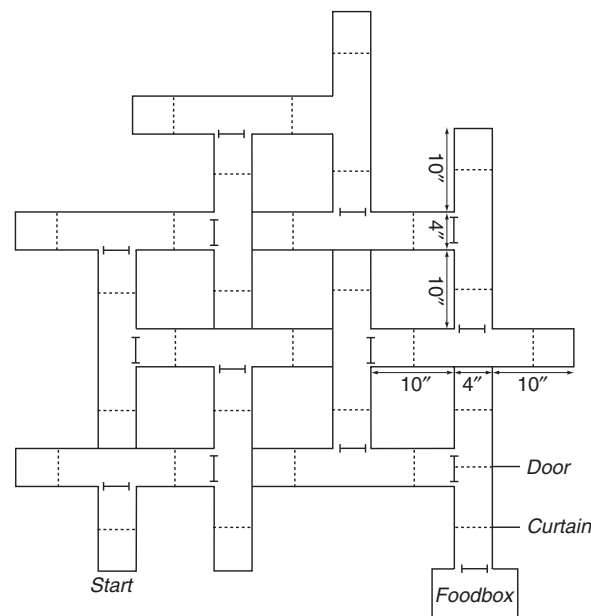
Perhaps the simplest way to demonstrate that the instrumental response is motivated by the expectation of the outcome is to change the value of the outcome after instrumental learning. The logic is straightforward; if desire for the outcome is motivating the instrumental response, then rendering the outcome less desirable should reduce the motivation to acquire it. Thus, reducing the value of the instrumental reinforcer after training should render it less desirable and weaken the instrumental response that had previously earned that reinforcer. Likewise, increasing the value of a reinforcer should increase the motivation to work toward obtaining that reinforcer.

Tolman and Honzik (1930) provided early empirical support for the role of the outcome expectancy in motivating the instrumental response (see also Tinklepaugh, 1928). Their study involved three groups of rats placed in a complex alley maze consisting of 15 T-mazes linked consecutively to each other (Fig. 5).

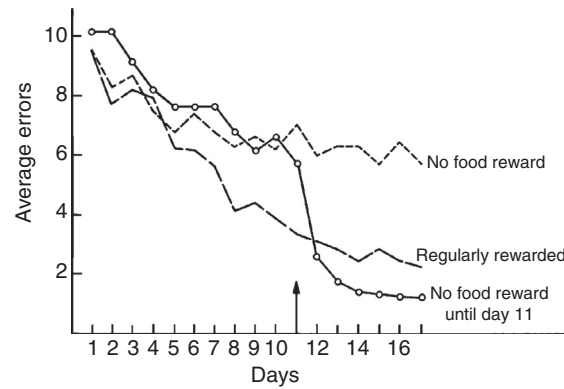
Each rat was placed in a start box at one end of the maze and removed from a goal box at the other end. Only one sequence of arms of the maze led from the start box to the goal box, and a number of other arms led to dead ends. Tolman and Honzik scored the number of dead end arms rats entered ("errors") during each trip from start to goal box. One group of rats was rewarded with food when they reached the goal box, and the number of errors decreased dramatically as a function of the number of trips they made (Fig. 6).

The second and third groups of rats received no food in the goal box for the first 10 trials in the maze. One of these groups, however, did find food in the goal box on the 11th and subsequent trials. Fig. 6 reveals that after finding food in the goal box for the first time on the 11th trial, the number of errors this group of rats made dramatically decreased on the 12th and subsequent trials. This result is perhaps the most famous demonstration of latent learning in the literature, showing that rats had learned to traverse the maze efficiently (i.e., with few turns into dead ends) even in the absence of explicit food reinforcement. In fact, performance was just as strong (if not stronger—reminiscent of a positive anticipatory contrast effect) in the group that was not rewarded during the first 10 trials as it was for the group for which food was available from the very beginning of training. More important for our purposes, this study also demonstrates that an increase in the motivational significance of a goal (in this case the contents of the goal box) has a dramatic and immediate effect on performance. The rats in the latent learning group must have changed their representation of the outcome on the 11th trial which motivated their performance on the 12th trial. More recent demonstrations of positive behavioral contrast effects like this provide further support for the role of outcome representations in mediating the strength of instrumental behavior (for a review, see Flaherty, 1996).

Tolman and Gleitman (1949) also showed that a downshift in the cues surrounding the reinforcer can affect instrumental performance. Rats first learned to find food at either end of a T-maze. The goal box at the end of one arm was always dark and that at the end of the other arm was always lit. Both goal boxes, however, always contained the same amount of food and thus had equal value. Following training, the rats were taken to a separate room that was dark in which they received footshocks. After this experience, the rats were returned to the T-maze. Not surprisingly, the rats avoided the dark goal box in favor of the lit goal box, despite both boxes



**Figure 5** Maze used by Tolman and Honzik (1930) to study latent learning in rats. From Tolman, E.C., 1948. Cognitive maps in rats and men. *Psychol. Rev.* 55, 189–208.



**Figure 6** Acquisition performance from Tolman and Honzik (1930). From Tolman, E.C., 1948. Cognitive maps in rats and men. *Psychol. Rev.* 55, 189–208.

containing the same amount of food and having no shocks. This preference suggests that the rats had encoded the illumination properties of the two goal boxes and avoided the dark box in which they expected the possibility of receiving another shock.

One problem with the above studies is that they involve a change in the quality of the reinforcer, such as from no food to food in the goal box or from food to food + shock in the dark part of the box. Perhaps the new outcome itself caused rapid changes in the instrumental response through the normal S–R mechanisms, rather than affecting the representation of the outcome mediating the response. That is, the added or altered outcome reinforces a new S–R association. This problem can be avoided by directly manipulating the value of the reinforcer itself without affecting its qualities or attributes. Adams and Dickinson (1981) demonstrated in an operant lever-pressing preparation that a downward shift in the value of the reinforcer immediately affects instrumental performance (but see Adams, 1980, 1982). Rats first learned to press a lever by reinforcing lever pressing with a food reward. After subjects had acquired the task, the reinforcer was devalued through pairings with a mild toxin that produced gastric malaise. After being paired with the toxin, rats were highly reluctant to consume the food, showing that they had acquired an aversion to it. When placed back into the conditioning chamber with the lever available that had earned the food during initial training, rats were now reluctant to press the lever as well. Importantly, the devaluation procedure was conducted away from the conditioning situation, which rules out effects of the manipulation on learning of the response and on context–illness associations. Moreover, the devaluation procedure had an immediate effect on instrumental responding when the rat was returned to the operant chamber.

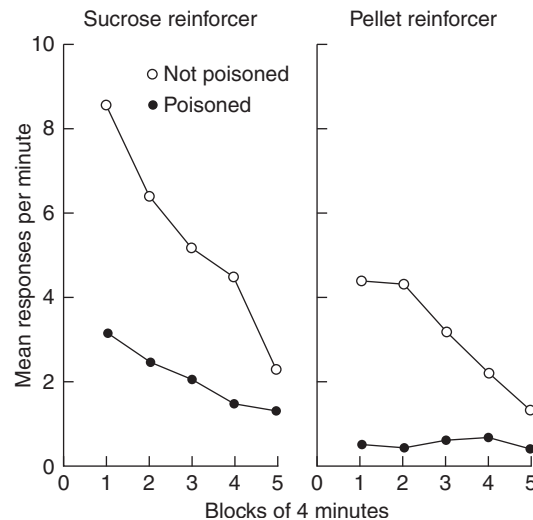
More compelling evidence for the role of outcome representations in the mediation of an instrumental response comes from work by Colwill and Rescorla (1985). They demonstrated the specificity of the devaluation manipulation to the instrumental response that had earned that outcome using a choice procedure. In their study, rats were trained on two action–outcome contingencies involving two actions (lever pressing and chain pulling) and two outcomes (a sugar solution and food pellets). For each subject, one of the actions was always reinforced with one of the outcomes (e.g., lever press → sugar solution), while the other action was always reinforced with the other outcome (e.g., chain pull → food pellet). Following instrumental training of the two action → outcome sequences, one of the outcomes was devalued through pairings with a mild toxin. The effect of this devaluation procedure was to depress instrumental responding of the action that had previously earned that outcome, but not of the action that had earned the other outcome (Fig. 7).

Similar results are obtained if the animal is satiated on one of the outcomes prior to the choice test session. Balleine and Dickinson (1998b) trained rats to make two instrumental responses (right and left lever presses), one for a salt-flavored and one for a lemon-flavored polycose solution. Following instrumental training, rats were given 1 h to feed freely on one of the foods immediately prior to an extinction choice test in which both responses were available (but no reinforcement was delivered during the test session). Response rates were significantly lower on the lever that had earned the food reward to which they had been satiated than on the lever that had earned the nonsated reward. The results of these experiments reveal the selective nature of the devaluation procedure and provide strong evidence that two separate R–O associations motivated instrumental responding. The selective nature of the devaluation treatments further shows that the qualitative features of each outcome were specifically associated with the instrumental response that earned that outcome. One important caveat needs to be mentioned. In most cases, the suppressive effects of outcome devaluation on instrumental responding require the subject to reexperience the devalued food during the extinction choice test (i.e., incentive learning; Balleine, 1992; Balleine and Dickinson, 1998a,b). Initial failures to find outcome devaluation effects on instrumental responding stemmed from failures to expose the subject to the devalued outcome after devaluation treatment (Adams, 1980, 1982).

#### 1.06.3.1.2 Manipulations of the Response–Outcome Contingency

A second line of evidence for R–O associations comes from manipulations of the R–O contingency. If an instrumental response is goal directed, then by definition it should be sensitive to the relation between the action and the outcome. If the outcome is made freely available, for example, then there is no need for the individual to go through the extra effort involved in making the





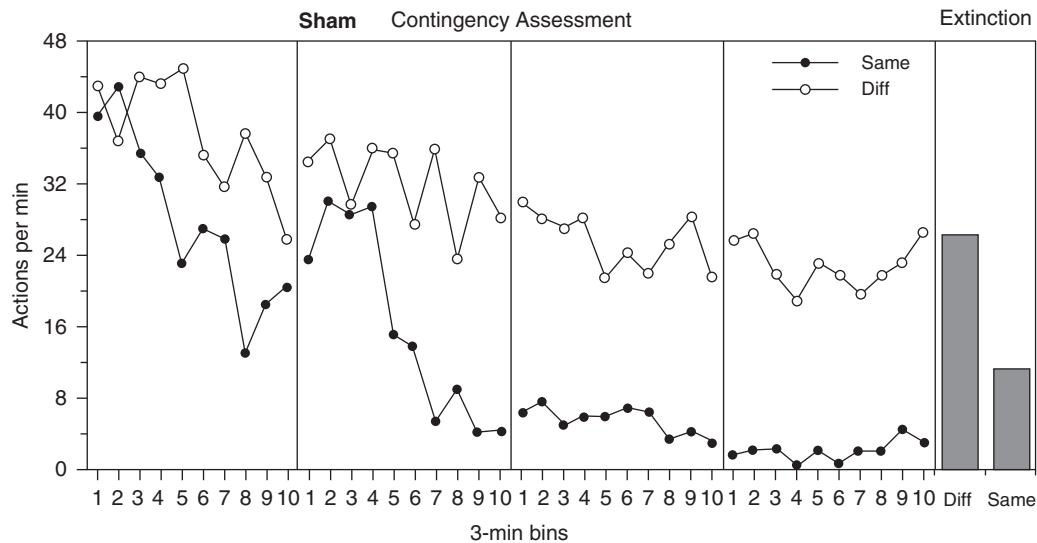
**Figure 7** Experiment 1 of Colwill and Rescorla (1985). Mean response rate on the levers that had earned the poisoned (filled symbols) and nonpoisoned (open symbols) actions, for both sucrose reinforcer (left panel) and pellet reinforcer (right panel). Reprinted from Colwill, R.M., Rescorla, R.A., 1985. Postconditioning devaluation of a reinforcer affects instrumental responding. *J. Exp. Psychol. Anim. Behav. Process.* 11, 120–132, with permission from R.A. Rescorla.

instrumental response to earn that outcome. Evidence showing the necessity of the R–O contingency on instrumental behavior comes from experiments using the omission procedure. In an omission procedure, a behavior is initially conditioned through pairings with reinforcement. After conditioning is established, the reinforcer is scheduled to be omitted if the subject makes a response. The reinforcer will be delivered, however, if the subject withholds responding. This procedure was developed for the purpose of dissociating instrumental from Pavlovian conditioned responding. The logic of the procedure is that Pavlovian CRs are not sensitive to the response–reinforcer contingency and thus should not be affected by the omission of the reinforcer; whereas instrumental responses which by definition are sensitive to their own consequences should be affected by the omission of the reinforcer. This procedure has been useful in discriminating acquired responses that are Pavlovian or instrumental in nature. For example, Holland (1979) showed that the acquisition of magazine approach during a tone that was paired with food developed despite the fact that food was withheld if the animal approached the food magazine during the tone. Although magazine approach appeared goal directed, it was actually shown to be insensitive to the negative response–reinforcer contingency. Rather, magazine approach appears to be a Pavlovian response (see Dickinson, 1988 for a fuller discussion of the omission procedure).

More recently, demonstrations that animals are sensitive to the R–O contingency have been pursued for the explicit aim of showing the goal directedness of instrumental behavior. A simple method that demonstrates the goal-directed nature of instrumental responding is to deliver free (i.e., noncontingent) reinforcers during an instrumental session. Delivery of noncontingent food pellets during a session in which a rat is engaged in pressing a lever for food suppresses instrumental lever pressing. Moreover, if a rat has two manipulanda available, each delivering a different outcome, the suppressive effects of noncontingent delivery of one outcome are selective to the manipulandum that earns that outcome (Colwill and Rescorla, 1986; Dickinson and Charnock, 1985; Dickinson and Mulatero, 1989; Hammond, 1980). For example, Colwill and Rescorla trained rats on two instrumental responses (lever press and chain pull), each for a particular outcome (sucrose solution and food pellet). After response rates had stabilized, one of the outcomes was freely delivered in addition to being earned by the response. The rate of responding on the manipulandum that earned the noncontingent outcome decreased, while the rate of responding on the other manipulandum was unaffected. In some cases, noncontingent presentations of one outcome can depress both responses that earn that outcome and those that earn a different outcome, but the response that earns the same outcome that is made freely available shows significantly greater suppression (Fig. 8; Balleine and Dickinson, 1998a).

### 1.06.3.1.3 Outcome Expectation Drives Response Variation

A third line of evidence has recently emerged that also suggests that the expectation of the outcome plays an important role in operant behavior. It has been well known since the middle of the last century that during the extinction of a previously reinforced operant, variability in the response dramatically increases (Antonitis, 1951; Balsam et al., 1998; Eckerman and Lanson, 1969; Frick and Miller, 1951; Herrick and Bromberger, 1965; Millenson and Hurwitz, 1961; Neuringer et al., 2001; Notterman, 1959; Stebbins and Lanson, 1962). This inverse relationship between reinforcement and response variation suggests (1) that during operant learning the subject acquires an expectation of the reinforcing outcome and (2) that response variation is inversely related to outcome expectation. Blaisdell has recently proposed a Law of Expect as a general principle of learning incorporating both postulates (1) and (2) above. The Law of Expect operates in conjunction with the Law of Effect to govern operant behavior (Blaisdell et al.,



**Figure 8** The mean number of actions (lever pressing and chain pulling) per minute during the four sessions of training under the noncontingent training and during the final extinction test (right-hand panel). The response rates are shown separately for the actions paired with the reward that was same as and different from the unpaired reward. Reprinted from Figure 1, Balleine, B.W., Dickinson, A., 1998. Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology* 37, 407–419, © 1998, with permission from Elsevier.

2016). Thus, when the expectation of reinforcement is high, operant behavior tends to be more stereotyped with low variation. When the expectation of reinforcement is low, however, such as during extinction, then behavior becomes much more variable.

Blaisdell and colleagues have developed procedures that provide a clear test of this account (Gharib et al., 2004; Stahlman and Blaisdell, 2011a; Stahlman et al., 2010a,b). Fig. 9 shows an example from Stahlman and Blaisdell (2011b) using operant screen pecking by pigeons in a touch screen–equipped operant chamber. Pigeons were presented with four types of trials in each session. On each trial, a discriminative stimulus (DS, a colored circle) was presented on the center of a touch screen. The pigeon had to complete a random-ratio (RR) 5 requirement of pecking to the DS to end the trial. Upon termination of a trial, the DS disappeared from the screen and an intertrial interval began. Of the four DSs, two were associated with a 12.5% probability of reinforcement, while the remaining two were associated with a 4.4% probability of reinforcement. If reinforcement was scheduled, it was delivered after the trial terminated.

For a small reinforcer delivered immediately upon trial termination, Stahlman and Blaisdell (2011b) replicated the greater response variation to the lower probability DS (4.4%) than to the higher probability DS (12.5%) originally reported by Stahlman et al. (2010a). This supports the Law of Expect, showing that the lower the expectation of food reinforcement (the outcome), the greater the variation in operant response (screen pecks). This was true for both the spatial location of screen pecks as shown in Fig. 9 as well as for variation in interpeak interval (not shown).

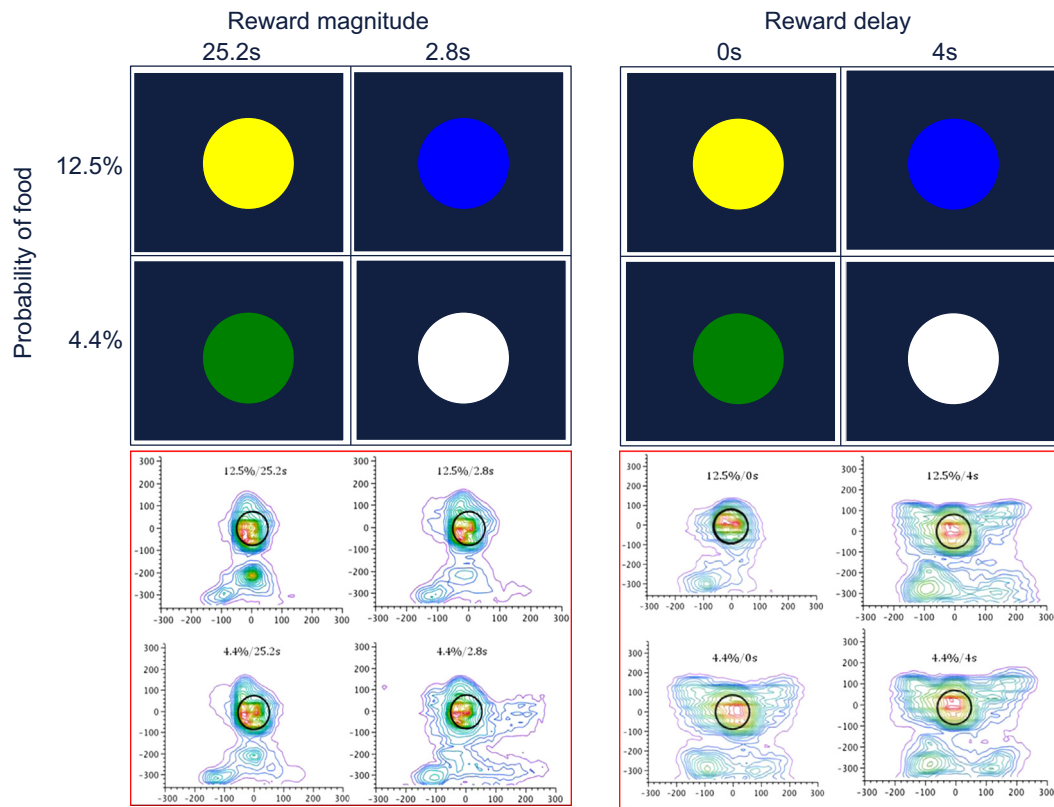
Stahlman and Blaisdell (2011b) also manipulated outcome expectation by systematically varying the magnitude of reinforcement (Fig. 9, left panels) or delay to reinforcement (Fig. 9, right panels). When the size of the reinforcer increased from 2.8-s access to food to 25.2-s access to food, thereby increasing the expectation of food reinforcement, response variation decreased to the smaller-probability DS (that signaled a 4.4% probability of food, Fig. 9, bottom left panel). Likewise, when the delivery of reinforcement was delayed until 4 s after the trial terminated, thereby decreasing the expectation of food reinforcement, response variation increased to the larger-probability DS (that signaled a 12.5% probability of food, Fig. 9, bottom right panel). Thus, any manipulation that reduced expectation of reinforcement resulted in an increase in response variation, whereas any manipulation that increased expectation of reinforcement reduced response variation.

### 1.06.3.2 Neurobiological Dissociations

The distinction between habitual and goal-directed behavior can also be made at the level of neural circuitry. Vertebrates and invertebrates both show parallel circuitry for reflexive behavior—including unconditional reflexes and species-typical fixed-action patterns on the one hand, and conditioned reflexes on the other—and voluntary behavior. For example, there are two circuits that mediate tail-flick escape behavior in the crayfish (*Procambarus clarkii*), one that mediates rapid and automatic escape responses and one that mediates slower and more flexibly controlled responses (Edwards et al., 1999; Wine and Krasne, 1972; see Fig. 10).

The rapid escape reflex in response to abrupt stimulation, such as a sharp tap to the side of the abdomen, is a fixed-action pattern that is mediated by medial giant (MG) command neurons. Intracellular recordings from the MGs detect electrical responses in as little as 10 ms after the tap stimulus is applied. The nongiant system, which is excited by gentle prodding and pinching, mediates longer-latency responses that are under a much greater degree of control by the animal than are the immediate escape behaviors. The





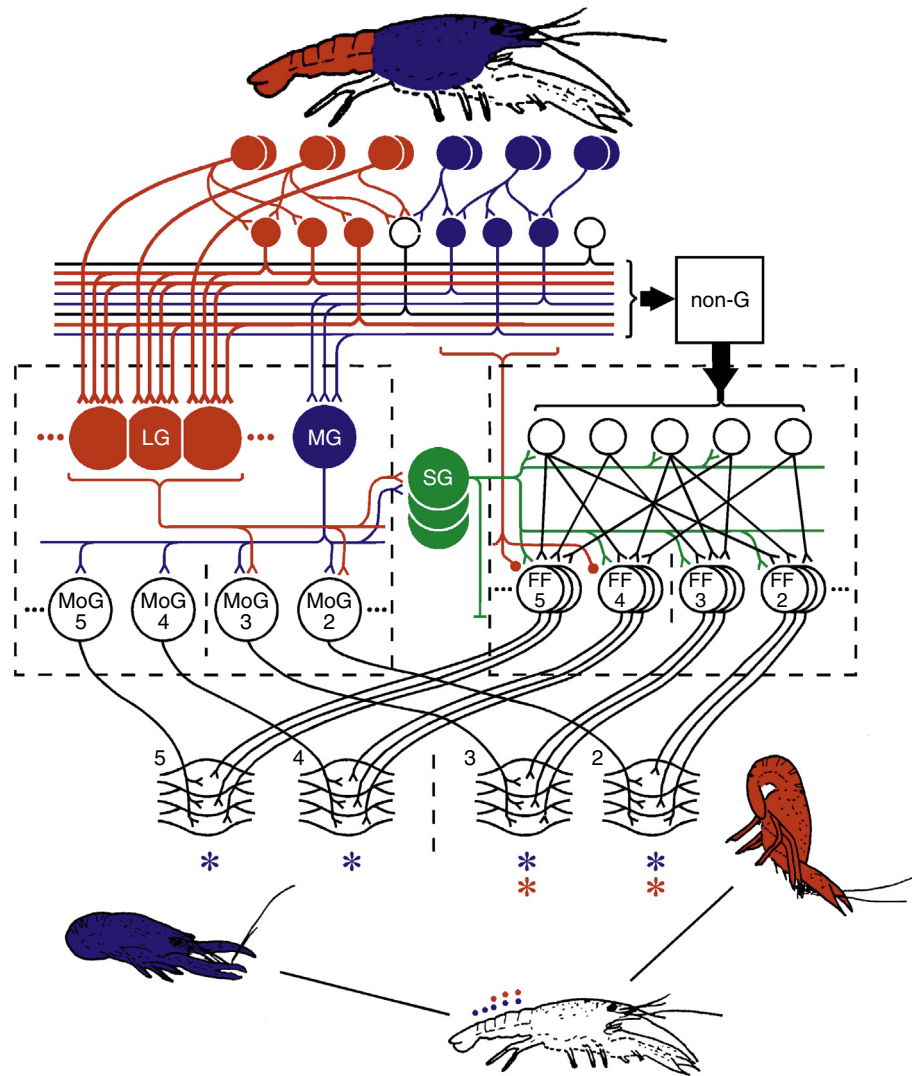
**Figure 9** Top panels: Diagrams of the  $2 \times 2$  within-subject design. Each of four discriminative stimuli was associated with one of two probabilities of reinforcement and one of two magnitudes of reinforcement (top left) or one of two delays to reinforcement (top right). Bottom panels: Nonparametric density plots illustrating the mean spatial location of pecks on the touch screen for each trial type, collapsed across subject, for manipulations of magnitude of reinforcement (bottom left) and delay to reinforcement (bottom right). The peck density plot models a smooth surface that describes how dense the data points are at each point in that surface and functions like a topographical map. The plot adds a set of contour lines showing the density in 5% intervals in which the densest areas are encompassed first. The *black circle* indicates the location and size of the stimulus disk, which is centered at (0,0). Units on both x- and y-axes are in pixels. Reprinted from Stahlman, W.D., Blaisdell, A.P., 2011. The modulation of response variation by the probability, magnitude, and delay of reinforcement. *Learn. Motiv.* 42, 221–236, with permission.

nongiant neural circuitry innervates and controls the same muscle systems as do the MGs, but is much more complex (Fig. 10) in both the interconnections and the number of layers between the sensory and motor neurons. Although much less is known about the functional control by the nongiant system, presumably it allows for a finer degree of control over the timing and direction of the movement and may even monitor actions as they are planned to allow for corrective feedback prior to execution of the action (see section [Making Things Happen](#)).

The distinction between the neural basis of the habit system and the voluntary or goal-directed action system can be made in vertebrates as well. There is not sufficient space here to adequately review the extensive literature on this dissociation, but it appears that in mammals at least S–R habit learning can be mediated at many locations within the nervous system, including the spinal cord (Chen and Wolpaw, 1995), the basal ganglia (White, 1989), and the striatum (Yin et al., 2004), while goal-directed R–O learning is mediated by cortical structures, such as the prefrontal area and the insular cortex (Balleine and Dickinson, 1998a). Similar pathways and circuits appear to be involved in modulating these two forms of instrumental control (Balleine and O'Doherty, 2010).

#### 1.06.4 Stimulus–Outcome Associations in Instrumental Discrimination Learning

In addition to R–O associations, evidence supports the role of S–O associations in controlling the instrumental response. S–O learning involves the formation of an association between the stimulus context in which the response is reinforced and the reinforcer itself. The stimulus context can either be diffuse situational cues, such as the puzzle box in Thorndike's experiments, or more discrete and punctate cues such as a light or a tone that is present during the time a response–reinforcement contingency is in effect. These cues are called discriminative stimuli (DSs), and they have been shown to enter into a direct association with the reinforcer, referred to as an S–O association (Colwill and Rescorla, 1988).



**Figure 10** Neural circuitry mediating the escape reflex and volitional tail flipping in the crayfish. Giant fiber (GF)–mediated reactions are portrayed in the drawings at the bottom of the left side: the red crayfish represents a lateral giant (LG) axon–mediated response and the blue crayfish represents a medial giant (MG) axon–mediated response. The segmental joints at which bending occurs to produce these reactions are indicated by *small colored circles* above the white crayfish. LG-associated elements and MG-associated elements are colored in red and blue, respectively. The sensory fields (mechanosensory for LG and mechanosensory and visual for MG) for the two types of GF-mediated reactions are indicated at the top of the figure. Circuitry for GF-mediated responses is shown on the left with primary afferents, sensory interneurons, LG and MG, and giant motoneurons (MoGs) arranged from top to bottom. The multisegmental nature of the LG, which is an electrically well-coupled chain of segmental neurons, each with its own dendrites, is indicated. *Colored asterisks* mark phasic flexor muscles of segments 2–5 that are used in each type of GF reaction. Circuitry for responses that do not use giant neurons (non-G responses) is shown on the right. A separate population of fast flexor (FF) motoneurons generates non-G responses; uncharted circuitry (*box* marked non-G) and a set of partially identified premotor interneurons (*open circles*) mediate between sensory neurons and FF motoneurons. The segmental giant (SG) neuron (green), with its blind-ending axon allows the LG and MG to recruit non-G motor and premotor units. LG neuron–associated sensory circuitry provides inhibitory input to caudal FFs (red) so that the SG will not cause bending at caudal joints during LG-type tail flips. *Curly brackets* show that multiple neurons of the population innervate the indicated target. Reprinted from Edwards, D.H., Heitler, W.J., Krasne, F.B., 1999. Fifty years of a command neuron: the neurobiology of escape behavior in the crayfish. *Trends Neurosci.* 22, 153–161, © 1999, with permission from Elsevier.

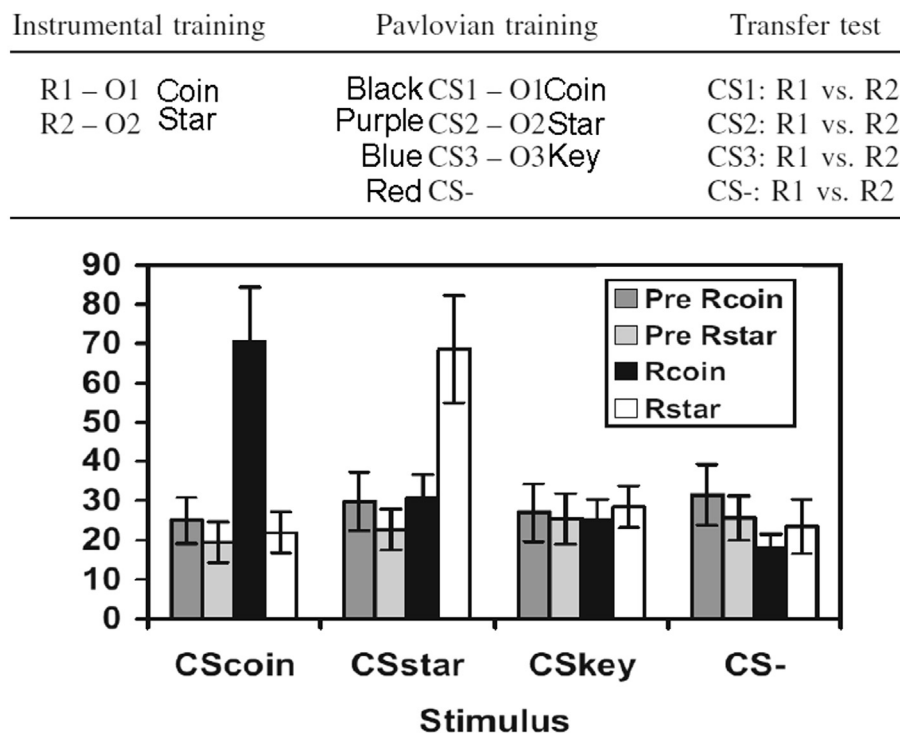
S–O associations became of central importance in two-process theories of instrumental learning early in the history of learning theory, starting with [Miller and Konorski \(1928\)](#); see historical review by [Rescorla and Solomon, 1968](#)). According to two-process theory, instrumental learning is governed by two processes, a Thorndikian process for the formation of S–R associations between the stimulus and response, and a Pavlovian process for the formation of S–O associations between the stimulus and reinforcer. The strongest empirical evidence for the role of S–O associations in instrumental learning comes from studies of Pavlovian-Instrumental Transfer (PIT).

The study of PIT involves first training subjects on a Pavlovian conditioned stimulus–unconditioned stimulus (CS–US) relationship, such as a tone paired with delivery of food. Next, subjects receive instrumental conditioning of an operant response, such as a lever press, using the food as a reinforcer. To test for PIT, the subject then receives a test session in which the operant is no longer reinforced (i.e., extinction). Shortly after the session has started and the rate of operant responding has subsided due to extinction, the CS is presented. If operant responding is invigorated during the presentation of the CS, this constitutes evidence for the role of Pavlovian associative processes in energizing or motivating operant responding (e.g., Kruse et al., 1983). The strongest evidence for the role of all three associative processes, S–R, S–O, and R–O, comes from studies using an experimental design involving two CSs, two instrumental responses, and two outcomes. For example, Nadler et al. (2011) trained human participants in an experiment using a simple computer task (Fig. 11, top panel). During the instrumental phase, participants learned to press one button on the computer keyboard (R1) to earn one outcome (O1, e.g., a picture of a coin on the computer screen) and a different button (R2) to earn a different outcome (O2, such as a picture of a star on the computer screen). Next, the keyboard was temporarily removed and participants received Pavlovian training with four different CSs that consisted of the computer screen turning a different color. CS1 (e.g., a black screen) was paired with O1; CS2 (e.g., a purple screen) was paired with O2. CS3 (e.g., a blue screen) was paired with a newly introduced outcome (O3, e.g., a picture of a key on the computer screen). Finally, CS4 was not paired with any outcome and thus served as a CS–. Finally, in the Transfer test, each participant was presented with the keyboard and allowed to freely press the two buttons (R1 and R2) both before and during the presentation of each of the CSs from the Pavlovian training phase.

The lower panel of Fig. 11 presents results. What is clearly apparent is that before any CS is presented, the rate of responding on each key is relatively low. When CS1 is presented, responses on R1 increase dramatically. Likewise, when CS2 is presented, responses on R2 increase dramatically. Neither the presentation of CS3 nor CS4 (the CS–) causes any increase in response on either R1 or R2. The increase in responding due to the presentation of CS1 and CS2 demonstrates operation of the S–O associations. The specificity of the effect of presenting each CS on an increase in the response that was reinforced with the same outcome as signaled by the CS provides evidence for the operation of individual R–O associations.

### 1.06.5 Stimulus–Stimulus Associations in Instrumental Discrimination Learning

In addition to entering into an association with the outcome, the DS can enter into an association with other stimuli present in the learning context. Early evidence for S–S associations came from studies of second-order conditioning by Pavlov (1927) and of

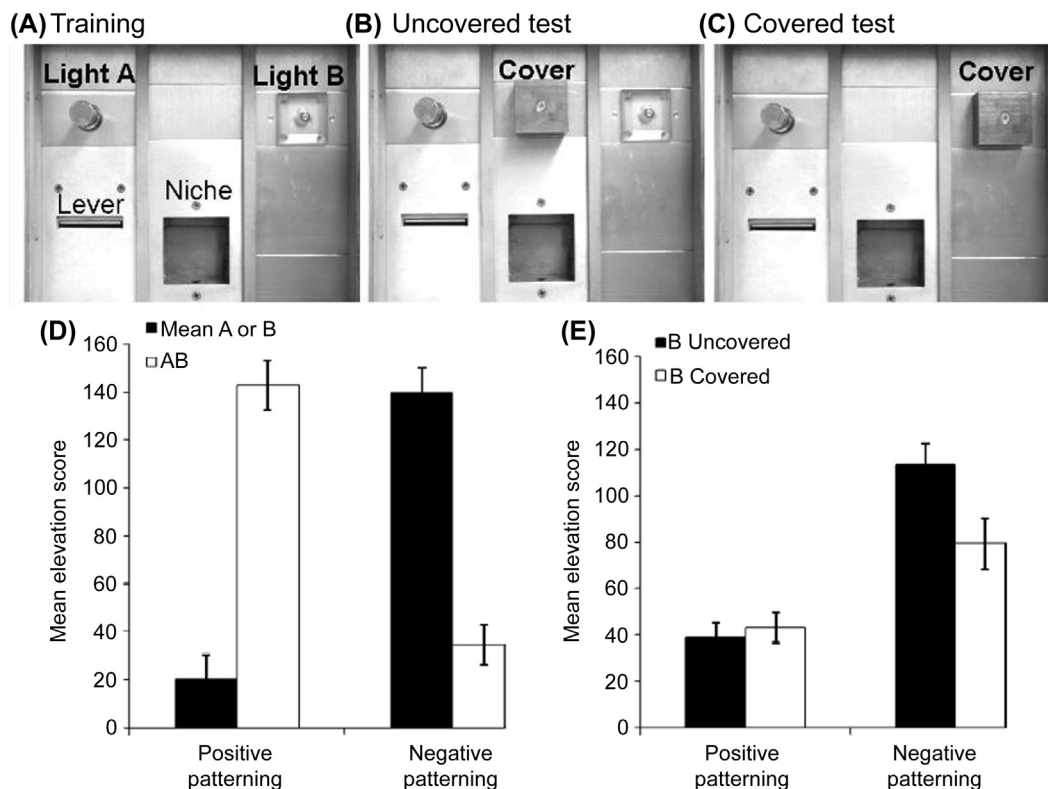


**Figure 11** Top panel: Experimental design used in the study of Pavlovian-Instrumental Transfer (PIT) in humans. R1 and R2 refer to different instrumental responses, O1, O2, and O3 refer to different reinforcing outcomes used during instrumental and Pavlovian training phases, and CS1, CS2, CS3, CS– refer to different conditioned stimuli (CSs) used during Pavlovian training. Bottom panel: Mean rate of instrumental responding during the PIT test. Data are shown separately in the prestimulus (Pre Rcoin and Pre Rstar) and stimulus periods (Rcoin and Rstar) for each test CS (CScoin, CSstar, CSkey, and CS–). Reprinted from Nadler, N., Delgado, M.R., Delamater, A.R., 2011. Pavlovian to instrumental transfer of control in a human learning task. *Emotion* 11 (5), 1112–1123, with permission from A.R. Delamater.

sensory preconditioning by Brogden (1939). Notably, these procedures involve the development of associations between Pavlovian cues. There is also evidence that S–S associations form directly between DSs during operant conditioning. The best evidence comes from studies of compound conditioning, in which multiple stimulus elements are presented in compound during learning. After learning, evidence for within-compound associations between stimulus elements comes from tests involving the omission of one or more elements. Such tests often reveal a generalization decrement in response strength. Furthermore, configural learning in vertebrates appears to involve the hippocampus (Sutherland and Rudy, 1989). Manipulations that inactivate the hippocampus disrupt behavior that depends on the formation of configural associations (e.g., within-compound associations) but not behavior supported only by elementary associations (e.g., S–O, S–R, and R–O).

Fast and Blaisdell (2011) provide an interesting demonstration of the special role of S–S associations in performance on an instrumental discrimination. Rats were trained on an instrumental discrimination with two lights (A and B) serving in either a positive patterning (A–, B–, and AB+; where the “+” and “–” refer to the presence and absence of food reinforcement, respectively) or negative patterning (A+, B+, and AB–) discrimination. Thus, discrimination was based on whether the DS+ was the presentation of the elements separately (as in negative patterning) or together in compound (as in positive patterning). The lower left panel of Fig. 12 shows that rats mastered each type of discrimination. Rats were then tested with only A lit while B remained unlit (explicitly absent) or occluded from view with an opaque shield (ambiguous whether B is lit or unlit, see Fig. 12, top panel). An interesting finding was that only rats that had been trained on the negative patterning discrimination responded differently when B was covered compared with when B was unlit (Fig. 12, bottom panel).

Why did rats that had been trained on a negative patterning discrimination respond less when tested on one light if the other light was covered? Fast and Blaisdell (2011) speculated that rats that had received negative patterning discrimination training had used a configural learning strategy to solve the discrimination. They did this by forming a within-compound association between lights A and B on compound trials. Thus, at test, the presentation of DS A could retrieve the representation of DS B by way of the within-compound association. When DS B is not covered at test, the rat could see that it was not present (B's bulb remained unlit) and thus the rat could discount its presence and respond as if it were an A-alone trial (which during training had led to



**Figure 12** Negative-patterning training is necessary for rats to infer the presence of a hidden cue. Rats learned to discriminate between illumination of one or two lights (A; lights A and B, counterbalanced), before being tested with only A lit while B either was uncovered and unlit (B) or covered by an opaque shield (C). (D) Mean elevation scores from the final session of pattern discrimination training in Experiment 1. Successful positive patterning (PP) discrimination is shown as higher response rates on compound (AB) trials than on elemental trials (A or B). Successful negative patterning (NP) discrimination is shown as higher response rates on elemental trials (A or B) than on compound (AB) trials. (E) shows mean elevation scores from trials with Cue A from the first test session. Subjects in Group NP lever pressed more during Cue A with B uncovered than when B was covered. Subjects in Group PP showed equally low rates of lever pressing during Cue A with B covered or uncovered. Error bars represent the standard errors of the means. Reproduced from Fast, C.D., Flesher, M.M., Nocera, N.A., Fanselow, M.S., Blaisdell, A.P., 2016. Learning history and cholinergic modulation in the dorsal hippocampus are necessary for rats to infer the status of a hidden event. *Hippocampus*. doi: 10.1002/hipo.22564.

reinforcement of lever pressing). When DS B was covered at test, however, the rat could not tell that it was absent. Thus, once the representation of B was associatively retrieved by DS A, it remained active. As a result, the rat acted as if it might be a compound trial, which had not been reinforced during training, and therefore the rat pressed the lever less. Subsequent work has confirmed an account of this phenomenon in terms of an associatively retrieved representation by way of the S–S association (Fast et al., 2016a), as well as the role of the hippocampus in maintaining the active representation of the hidden event (i.e., DS B when covered; Fast et al., 2016b; see also Blaisdell et al., 2009; Waldmann et al., 2012).

### 1.06.6 Agency

#### 1.06.6.1 Intentional Psychology: Beliefs and Desires

Now that we have established the veracity of the goal directedness of some acquired behaviors, we can now speak with some assurance about the role of beliefs and desires in an intentional psychology (Dickinson, 1988). Intentionality is notoriously difficult to establish in a system because so much behavior can be described using intentional language. We readily slip into intentional language to describe the behavior of our pets (Rover is barking to warn the stranger at the doorstep), computers (My laptop is searching its memory for the file), and the weather (Political leaders escaped Katrina's wrath). The language of intentional folk psychology provides a convenient shorthand to describe and explain behavior, even when we fully recognize that the behavior itself is not intentional at all (Dennett, 1989). Learning theorists themselves find it difficult not to slip into intentional language when describing the behavior of their subjects. In fact, we frequently apologize to our peers for slipping into teleological or anthropomorphic language when we catch ourselves doing so or diffuse the affront by placing the offending material in scare quotes. The rampant misuse of intentional language, however, is for the most part benign and should not devalue the scientific investigation of behavior that is truly intentional in nature. In fact, describing nonintentional behavior using intentional language may assist humankind's ability to predict and control nonintentional systems (Dennett, 1989).

Given the evidence for goal-directed, intentional behavior in the Animal Kingdom, what is its function? That is, why would it have evolved? Could not a creature that lacked goal-directed behavior function just as well in its world as one that had it? Perhaps not. Explicit representation of a goal may serve to motivate behavior to bring the animal into contact with the goal. Furthermore, an animal that believes that an action produces (or prevents) a goal and that desires (or dislikes) that goal can be more flexible in when and how it goes about seeking and obtaining (or avoiding) the goal. For example, an omnivorous species—such as a rat, a pig, or a human (or a honeybee)—must learn which foods are good to eat and which should be avoided. A food previously discovered to be safe for consumption, however, may become spoiled or otherwise unpalatable. The animal that is capable of learning about the devalued food should be able to refrain from seeking out and consuming that food more rapidly than an animal that depended on trial-and-error learning alone. A rat is much more likely to live after one poisoning event than after many. Likewise, an animal that learns of another source where food can be obtained with much reduced effort should be able to immediately curtail exertions to acquire that food from its previous source. Such an animal can plan for the future. So far, these suppositions are no better than Kipling's (1912) *Just So Stories*. There may, however, be a more important and defensible role for intentional behavior that led to its evolution. This has to do with the concept of agency – a term that often arises in discussions of goal-directed action and intentional behavior.

#### 1.06.6.2 Animals as Free Agents

Leslie (1995) defines an agent as an object that has three properties that distinguishes it from other physical objects: mechanical, actional, and cognitive properties. The mechanical property that distinguishes agents from other types of objects is that they have an internal and renewable source of energy (or FORCE in Leslie's terminology) that allows them to *cause* things to happen in the world without themselves having to rely on external sources of energy and force (although we must acknowledge the fact that even agents must refuel). Premack (1990) has made the same distinction invoking the term "self-propelledness." This property allows the agent to be a source of causation (though perhaps not the ultimate source according to adherents of philosophical determinism). To be classified as agents, objects also need to exhibit actional properties, which consist of the object's ability to act and react to events or circumstances in the world that are spatially and/or temporally distant. This property is what characterizes the agent's behavior as goal directed or intentional – at least by appearance. Note, this property of agenthood can be ascribed to an object which only appears to have intentional, purposeful, goal-directed behavior, even if the internal control mechanism itself is not goal directed. For example, a moth's suicidal plunge into a burning flame appears goal directed (the moth sought the flame) even though the control mechanism is a simple, innate (i.e., "blind") phototactic reflex. The third property of agency is the cognitive property. An agent's behavior is determined by its beliefs about the world (propositional knowledge). Beliefs are not only about static, coherent states of the world (semantic knowledge) but, more importantly, are about the causal texture of the world (Blaisdell, 2009; Gallistel et al., 2001; Racey et al., 2011). To believe that A causes B asserts the belief that certain values of A (a1, a2, a3...) are determinants of certain values of B (b1, b2, b3...) and that changes in the state of A (e.g., a1 → a2) should bring about a complementary change in the state of B (i.e., b1 → b2) (see Woodward, 2003 for a discussion of causal explanation).



## 1.06.7 Interventions and Causal Reasoning

### 1.06.7.1 Making Things Happen

What advantages does an agent gain with its ability to represent causal relationships in this manner and hold desires for particular outcomes? One obvious advantage is that such an agent could test its belief system—that is, to fact check. More importantly, an agent could check facts systematically (cf. [Dennett's \(1995\)](#) Popperian creatures) rather than through blind or random trial and error (cf., [Dennett's Skinnerian creatures](#)). The intersection of causal beliefs with desires for goals provides a creature with the ability to manipulate its environment to achieve its goals. This feature of agency provides a creature with a powerful tool—instrumental manipulation of its world. A purely Pavlovian creature can merely passively learn about the causal texture of the world through observation; such a creature is stuck in its world and can merely predict effects based on their cues and respond in anticipation. A creature with both Pavlovian and Instrumental learning processes available to it could both passively learn about the causal texture of the world through observation and actively manipulate the world to directly discover its causal texture. Goal-directed instrumental learning allows the agent to actively explore its world through direct intervention. This exploration will uncover many cause–effect relationships that would have remained hidden to a purely passive observer.

Another important feature of agency is that goal-directed behaviors impart a sense of agency to the individual's intentional behaviors. That is, in contrast to reflexive habits, goal-directed behavior is accompanied by unique sensory feedback that is responsible for the sense that “I” (the agent) control events in the outside world ([Haggard, 2005](#)). This distinction is supported by experimental ([Libet, 1985](#)) and neurobiological ([Sirigu et al., 2004](#)) evidence. These internal sensory markers of intentional behavior should be important, if not critical, to the ability to reason from causal interventions (see discussion below). That is, without the ability to distinguish effects resulting from self-generated, intentional actions from effects resulting from other causal sources (including the agent's nonvoluntary or reflexive behavior), the agent would be incapable of interventional reasoning. This suggestion is supported by imaging studies that use transcranial magnetic stimulation to temporarily inactivate a particular brain area. These studies suggest that the presupplementary motor area—which tracks the neural pathways responsible for intentional action—acts as an internal monitor of intentional action ([Haggard and Clark, 2003](#); [Haggard et al., 2002](#)). A similar role has been found for the cerebellum ([Schmahmann, 1991](#); [Vilis and Hore, 1980](#)). This predicting signal allows the agent to correct errors in the execution of the action before the action itself occurs. Moreover, this predictive signal can be used to enhance the perception of non-self-generated sensations. For example, the tactile sensation that results from one finger touching the adjacent one is perceived as weaker than the same stimulus imposed from an external source. Attenuated self-generated sensations can more readily be ignored, freeing up attentional resources to focus on the external world. This attenuation of self-generated sensory experience may result from a predictive process that anticipates self-generated sensations or a postdictive process that judges the source of the perception to be self-generated or externally generated after the sensation has been experienced.

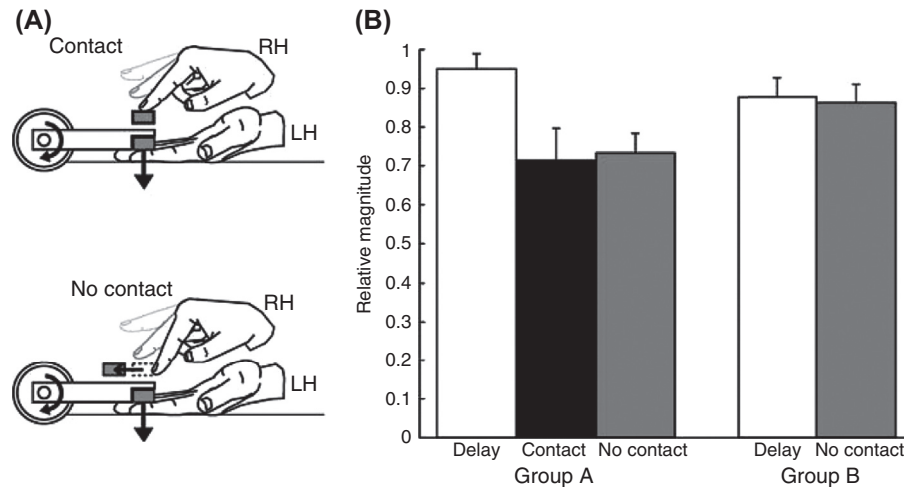
One study suggests that the process is predictive ([Bays et al., 2006](#)). Participants were required to judge whether a second (comparison) tap was stronger or weaker than a first (sample) tap to their left index finger. The first tap was always the same magnitude, while the strength of the second tap was varied across trials.

On most trials, the second tap followed the first by a short interval and was produced by the subject using tapping with their right hand onto a button positioned above the left-hand index finger (contact trials, [Fig. 13A](#), top panel). This produced a sensation of tapping on one's finger through a solid object. Occasionally, subjects would receive a trial on which the second tap was artificially delayed by 500 ms after they tapped the button with their right hand. On delay trials, the second tap was perceived as being much stronger than a tap of the same magnitude on contact trials ([Fig. 13B](#), Group A). The increased magnitude of the perceived tap on delay trials was not due, however, to the absence of a right index finger press at the time the tap was felt. Of primary importance, however, were the occasional no-contact trials that were conducted to test whether subjects were predicting a self-generated sensation when they attempted to press the button. On no-contact trials, the button was removed so that when the subject attempted to make a button press, they failed to actually press the button ([Fig. 13A](#), bottom panel). Despite the absence of a button press, the subjects reported these no-contact taps to be much weaker than the taps felt on delay trials. In fact, no-contact taps were perceived as being the same magnitude as contact taps ([Fig. 13B](#), Group A). These results show that the subjects anticipated the sensation of a self-generated tap, even when they were prevented from physically engaging the tap. It therefore appears that the process that monitors intentional actions is predictive and attempts to attenuate sensations arising from self-generated actions. This in turn can enhance the perception of sensations caused by an external source.

### 1.06.7.2 Seeing Versus Doing

A burgeoning literature exploring the theoretical mechanisms of causal interventions is developing in the fields of computer science, statistical theory, philosophy, and psychology ([Blaisdell and Waldmann, 2012](#); [Pearl, 2000](#); [Spirtes et al., 1993](#); [Sloman, 2005](#); [Steyvers et al., 2003](#); [Waldmann and Hagmayer, 2005](#); [Waldmann et al., 2006, 2008](#); [Woodward, 2003](#)). A resulting achievement is a clear and precise language, taxonomy, and formalization of the difference between observing the cause–effect relationships among a set of variables (Seeing) and intervening on one variable to determine its causal status in relation to the other variables (Doing). These analyses describe the special status interventions have on our (human) ability to determine cause–effect relationships in the world. “If I flick this switch, the light turns on. If I don't flick the switch, the light remains dark.” Such a simple cause–effect relationship can only be determined through intervention. (Note, the agent does not have to actively intervene, the agent can merely observe another agent intervening or observe a fortuitous intervention, such as a book falls off the shelf and accidentally





**Figure 13** (A) Schematic of the apparatus and task used by Bays et al. (2006). On contact trials (top), in response to an auditory go signal, participants produced a brief force pulse with their right index finger on a force sensor fixed above their left index finger. A similar force pulse was delivered to the left index finger by a torque motor. On no-contact trials (bottom), the force sensor was moved at the start of the trial so that participants made a tapping movement with their right index finger but did not make contact. (B) Mean relative magnitude of the comparison tap to the test tap at the point of perceptual equality as a function of trial type and participant group. Error bars represent 61 SE. Reprinted from Bays, P.M., Flanagan, J.R., Wolpert, D.M., 2006. Attenuation of self-generated tactile sensations is predictive, not postdictive. *PLoS Biol.* 4, e28, © 2006.

flicks the switch on its journey to the floor. Theoretical treatment of causal interventions does not treat these scenarios as being different in any significant way [Tomasello and Call, 1997].) The ability to reason about cause–effect relationships through the intervention on a single variable is the basis for the scientific method, which gives humankind an incredible analytical power over the world. “If I put reagent X into a beaker filled with reagent Y, the mixture ignites, otherwise the mixture remains inert.” “If I look at a blue-filled circle for 60 seconds, I then see a yellow, circular after-image when I look at a white wall.”

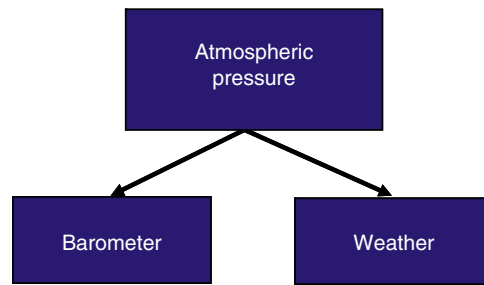
A simple example will serve to clarify the fundamental difference between Seeing and Doing, and the powerful role interventions (Doing) play in causal reasoning processes. Consider the workings of a barometer (Fig. 14).

The barometer’s reading may vary upward or downward, and this variation correlates strongly with changes in the weather. If we observe an increase in the barometer’s reading, we can expect to see sunny skies when we look out the window. If we observe a decrease in the barometer’s reading, however, we can expect to see rain clouds gathering. This tight correlation between the barometer and the weather is not due to a direct causal relationship between the two events. Rather, changes in both the reading of the barometer and in the weather are caused by a third event—changes in atmospheric pressure. If we understand the actual causal relationship among these three variables, then we would predict that tampering with the barometer (an intervention) should not affect the weather (or air pressure for that matter; Waldmann and Hagmayer, 2005). Imagine a child observing a barometer for the first time. Without prior schooling on its operation, he or she might at first entertain the notion—due to temporal priority—that changes in the barometer’s reading cause changes in the weather. If he or she set the barometer to a higher reading, however, he or she would soon discover that he or she did not bring about sunny skies. Through intervention on the barometer he or she was able to test his or her hypothesis and discover a more accurate underlying causal structure. He or she would never have been able to discover the underlying causal relationship without access to interventional knowledge (his or her own or someone else’s). Gopnik et al. (2004) provide another telling example.

...smoking is correlated both with having yellow fingers and with getting cancer, so having yellow fingers is correlated with getting cancer, but cleaning your hands will not keep you from getting cancer, and quitting smoking will. Knowing the right causal structure may not be essential for predicting one thing from another, but it is essential for predicting the effects of interventions that deliberately manipulate events.

Gopnik et al. (2004, p. 8).

Goal-directed behavior provides an important foundation for interventional reasoning. The ability to use interventions to examine causal relationships within a system is predicated on three assumptions (adapted from Gopnik and Schulz, 2004): (1) that interventions are exogenous to the system being studied and not caused by other variables within the system, (2) the intervention directly fixes one variable within the system to a specific value (e.g., a switch is moved from OFF to ON), and (3) the intervention does not affect the values of other variables within the system except through its influence on the variable that is the target of the intervention. The second and third assumptions necessarily depend on goal-directed behavior. If an individual cannot represent the contingent relationship between their actions and goals, then they would not profitably use their actions—via the effects their actions have on the world—to investigate cause–effect relationships.

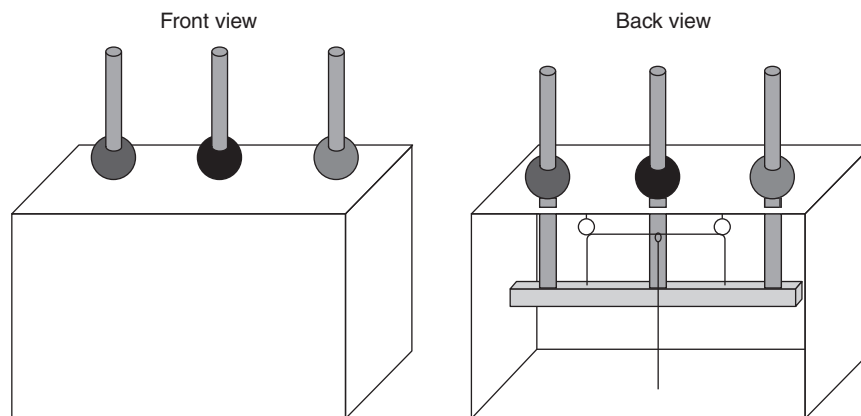


**Figure 14** Causal model for how a barometer and the weather are both effects of a common cause, atmospheric pressure. Arrows indicate causal relationship directed from the cause to the effect.

#### 1.06.7.2.1 Seeing Versus Doing in Children

Children, it turns out, do appreciate the special role of interventions in diagnosing cause–effect relationships from an early age (Gopnik et al., 2004). In one experiment, 4-year-old children were shown two or three colored rubber balls attached to sticks (hereafter referred to as “puppets”) and placed on the top of a box. The children observed the puppets move up and down simultaneously. The puppets could be attached to each other—although the children could not see the attaching mechanisms—so that the experimenter could arrange for the simultaneous movement of the puppets (Fig. 15). This enabled the children to observe the correlation of movement without observing the intervention by the experimenter. In other conditions, the experimenter could move one puppet at a time, independently of the others. In the pretraining phase of the experiment, the children watched the puppets move together simultaneously. They were then told that one of the puppets was “special” in that it always made the other puppet(s) move.

The children were then presented with two types of tasks. In the common effects task, the children observed two puppets (X and Y) move and then stop together simultaneously a number of times, followed by a demonstration that one of the puppets (Y) could be moved without affecting the movement of the other puppet (X). This was accomplished by the experimenter visibly moving puppet Y by moving the top of the stick to which it was attached. Finally, the children watched both puppets move together again by the experimenter invisibly moving the connecting mechanism behind the box. The children were asked “Which is the special puppet?” The children chose puppet X a majority (78%) of the time. This suggests that they drew the “correct” causal interpretation of the relationship between the puppets, that is, that movement of Y is a common effect of both the experimenter’s intervention (I) and of the movement of X. I use quotes because the movement of puppet Y was always due to intervention by the experimenter, however, the children could not see this intervention when it was hidden by the box. Thus, given their sensory data, they observed two conditions: puppets X and Y are moving without an apparent intervention by the experimenter, and puppet Y is moving in the absence of the movement of puppet X when the experimenter intervened on it. The children’s selection of puppet X as special indicates that they represented the causal relationships in the following way:  $I \rightarrow Y \leftarrow X$ , where the letters represent the events (I = intervention, Y and X = puppets) and the arrows indicate causal directionality. In causal graphs, arrows always point from the cause to the effect. Thus, in the causal graph the children inferred, Y was a common effect of both intervention by the experimenter (I) and of the movement of puppet X.



**Figure 15** The puppet machine used by Gopnik et al. (2004). Reprinted from Gopnik, A., Glymour, C., Sobel, D.M., Schulz, L.E., Kushnir, T., Danks, D., 2004. A theory of causal learning in children: causal maps and bayes nets. *Psychol. Rev.* 111, 3–32, with permission from A. Gopnik.

The second task involved a common cause relationship among events. In this task, children observed three new colored puppets (X, Y, and Z) move and stop together a number of times. Then they observed the experimenter intervene to move Y independently of the other two puppets and then intervene to move Z independently of the other two puppets. The causal relationships the children should have drawn from this pattern of observations can be graphed as follows:  $I \rightarrow Z \leftarrow X \rightarrow Y \leftarrow I$ . This causal graph follows from the observations that puppets Z and Y could be moved independently through intervention (I) or moved simultaneously with puppet X. Thus, the children should have inferred that X was a common cause (i.e., a special puppet) of the movements of the other two puppets. This is exactly what the children reported, choosing X as the special puppet 84% of the time.

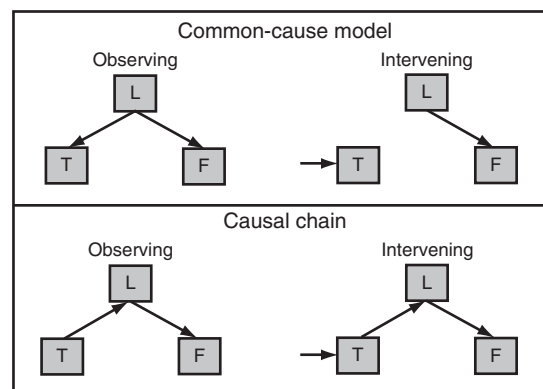
These experiments reveal how important information from interventions is for young children to reason about the causal texture of their world. If the children only had access to observations of the puppets moving together, then it would have been completely ambiguous which puppet or puppets caused the others to move, if indeed any of them were causally related to the others. It would also have been just as likely that some other, hidden force caused all of the puppets to move together (cf. changes in a barometer and the weather being driven by the unobservable air pressure). By observing that direct interventions on some of the puppets caused only those puppets to move, allowing the children to select out the unlikely causal relationships among them and zero in on the most likely interpretation. That is, interventions allowed them to test their hypotheses.

### 1.06.7.2.2 Seeing Versus Doing in Rats

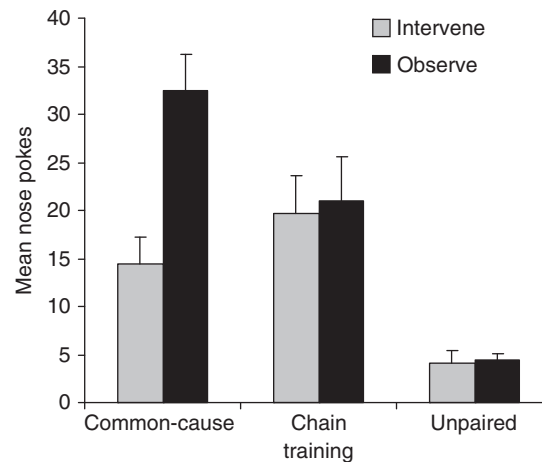
If children can reason about cause–effect relationships from interventions, what about other animals? What is the evidence that nonhuman animals can reason about their causal interventions? Blaisdell et al. (2006) investigated this question in rats using conventional conditioning procedures. In a training phase, rats observed a light followed by a tone ( $L \rightarrow T$ ) in one session and a light followed by food ( $L \rightarrow F$ ) in a second session (Blaisdell et al., Experiment 2).

One possible causal representation the rats could derive from this observational learning is that the light was a common cause of both the tone and food. No levers were present during the training phase, but during the test phase a lever was inserted into the conditioning chamber for the first time. Note that the rats had never seen this lever before and had certainly not received any training to press the lever. Nevertheless, the following contingency was put in place for half of the rats in the study: if the rat pressed the lever, the tone would come on. In contrast to this Intervention condition, the remaining rats were allocated to an Observation condition. Rats in the Observation condition also had the lever available, but pressing the lever had no effect. That is, the rats in the Observation condition had an inactive lever. The Intervention condition allowed Blaisdell et al. to assess what the rat inferred when it intervened via the lever press to turn on the tone. The Observation condition measured whether the rats expected food when they heard the tone. If rats had formed the causal model  $T \leftarrow L \rightarrow F$ , then by observing the tone (Group Observe) they should diagnostically predict the light and thus also predict that food should be available (Fig. 16, top panel, left-hand graph). By intervening on the tone, however, the rats should infer that they—and not the light—had caused the tone. Therefore, they should neither predict the light nor the food (Fig. 16, top panel, right-hand graph). This pattern of data was exactly what Blaisdell et al. observed (Fig. 17, left-hand bars).

If the rats were reasoning correctly about their interventions, then lever pressing should not invariably disrupt all causal relationships between the tone and other events, but only between other causes of the tone. For example, interventions on the tone should not disrupt expectations of the tone's effects. To show this, Blaisdell et al. (2006) tested another group of rats that had similar training as described earlier but for one key difference. Rather than receiving Light  $\rightarrow$  Tone pairings in the first phase of training, they received Tone  $\rightarrow$  Light pairings. This treatment, when combined with the following phase of Light  $\rightarrow$  Food pairings, should have taught the rats a Tone  $\rightarrow$  Light  $\rightarrow$  Food causal chain (Fig. 16, bottom panels). Again the rats were divided into two test groups after receiving causal-chain training, with half the rats receiving the Intervention condition (lever pressing turned on the tone) and the remaining rats receiving the Observation condition (tones were presented independent of lever pressing). If rats had learned to treat the light as an effect of the tone, then interventions on the tone at test should still activate an expectation of the light which



**Figure 16** Causal models of the relationships among events in Experiment 2 of Blaisdell et al. (2006). Reprinted from Blaisdell, A.P., Sawa, K., Leising, K.J., Waldmann, M.R., 2006. Causal reasoning in rats. *Science* 311 (5763), 1020–1022, © 2006.



**Figure 17** Data from [Blaisdell et al.'s \(2006\)](#) Experiments 2a and 2b. Adapted from Blaisdell, A.P., Sawa, K., Leising, K.J., Waldmann, M.R., 2006. Causal reasoning in rats. *Science* 311 (5763), 1020–1022, © 2006. See text.

should then generate the expectation of food ([Fig. 16](#), bottom panel, right-hand graph). Thus, equivalent amounts of magazine entries should be observed in both the Intervention and Observation conditions. This is exactly what was observed ([Fig. 17](#), central bars). It was furthermore established that magazine entries evoked by the tone depended on the tone having an indirect causal relationship to the food. This was shown through the inclusion of a third group of rats that had received unpaired presentations of the tone and light, so that the two could not enter into any kind of causal relationship, which prevented the tone from eliciting magazine entries at test ([Fig. 17](#), right-hand bars).

These results show that even rats can reason in a sophisticated manner about their goal-directed interventions by using them to infer the causal structure of the world. I am pitching this description of reasoning processes in rats at the computational level of analysis ([Marr, 1982](#)), which is described in terms of representations. The distinction should be made between a causal representation and the learning process through which the representation is acquired ([Dickinson, 1980](#); [Heyes and Dickinson, 1990](#)). The learning process might take the form of an entirely algorithmic level associative process. The rats for which a lever press produced a tone did not expect food as strongly as did rats that received tones unconnected to their lever pressing behavior. A plausible explanation for this difference at the algorithmic level of description lies in the rat's knowledge about what instrumental actions do. That is, the rats in our study (and rats in general) have had lots of experience learning about the effects of their actions on the world. This learning starts in infancy (if not earlier) and continues throughout the lifetime of the rat. In fact, both 1-day-old rat pups ([Johanson and Hall, 1979](#)) and newborn human infants ([DeCasper and Fifer, 1980](#)) can learn instrumental responses. Thus, throughout their lifetimes, rats acquire a large number of action–outcome associations. This learning is typically accompanied by unique feedback cues associated with self-generated, volitional behavior (see section [Making Things Happen](#)). These feedback cues are presumably present while the rat is engaged in lever pressing in the study phase of the test by [Blaisdell et al. \(2006\)](#). If the associations between these feedback cues and prior outcomes of self-generated, goal-directed action generalize to the current test situation, then the rats that experience the tone following their lever press should discriminate the tone as being the outcome of a self-generated cause rather than the outcome of a non–self-generated cause. Therefore, the tone in the intervention condition should be treated as being caused by the rat itself, and not by the light. The light should not be expected (at a higher rate than its baseline rate of occurrence) and hence neither should the food. In other words, there are strong grounds for believing that rats can distinguish between self-generated and non–self-generated events, allowing them to learn that they can produce effects through their actions. In support of this, [Leising et al. \(2008\)](#) found evidence that when followed by the tone, a rat's action on the lever was treated as a causal intervention, while an exogenous cue was not. Rats that received a tone when they pressed the lever may have generalized from their vast instrumental experience and treat the tone as the effect of their own action. This is a likely candidate psychological mechanism for reasoning about causal interventions in both nonhuman and human animals.

The analysis above suggests that, while the rats in the study by [Blaisdell et al. \(2006\)](#) used causal beliefs or models to reason about the source cause of the tone at test, these beliefs or models could have been acquired through conventional associative learning processes ([Dickinson, 1980](#)). Dickinson has argued that, in both animals and humans, associative learning processes can support the acquisition of causal beliefs (as distinct from associations) that then control performance through a process of practical inference (see also [Heyes and Dickinson, 1990](#)). Of course, humans also have other routes to causal knowledge available, such as verbal and instructional, but the associative (statistical) process is clearly of great importance to humans as it is to other species.<sup>1</sup>

<sup>1</sup>I am grateful to Tony Dickinson who is the source of most of the ideas presented in the preceding two paragraphs. His discussions with me have been immensely helpful.

Nevertheless, there are alternative frameworks to the associative one by which to conceptualize acquisition of causal knowledge (e.g., Cheng, 1997; Waldmann et al., 2008).

### 1.06.8 Tool Use: From Crow to Cro-Magnon

The intentional, goal-directed dissection of cause–effect relationships in the world depends on two abilities: (1) the ability to discriminate self-generated intentional acts from those elicited reflexively by the environment, and (2) the ability to discriminate changes in the environment due to one's own actions from changes produced by other means (e.g., environmentally caused). Hence, it is only through the intentional and systematic probing and manipulation of an object that one can build a representation of its causal features. Our knowledge of everyday folk physics—that unsupported objects fall toward the ground, that pushing or pulling on an object usually moves it in the direction of the push or pull, etc.—derives from our vast lifetime experience of manipulating objects and observing the dynamic world (Gopnik et al., 2004). The ability to construct causal maps of the world through observation and planned intervention (what Gopnik et al. refer to as egocentric causal navigation) is not confined only to the human animal, but is likely present in many other species, as well. I already presented above evidence that rats in a conventional laboratory setting are capable of deriving causal inferences from observational and interventional data. Below I review some of the recent literature investigating what type of causal understanding underlies tool use in nonhuman primates and in Corvids – a family of birds.

#### 1.06.8.1 Tool Use in Primates

Tool use provides an interesting case in which to study causal knowledge and reasoning. Tools are objects with properties and affordances that convey functional value to achieving a particular goal. For example, chimpanzees (*Pan troglodytes*) in West Africa learn to use hammer-anvil stone tools to crack open palm nuts and extract the meat inside (Fig. 18).

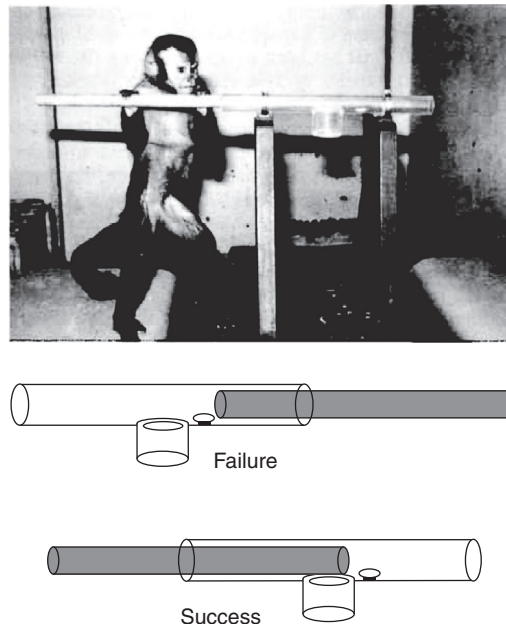
This behavior is transmitted culturally from one generation to the next. The full act, which takes many years to master, involves placing a round palm nut on a large, flat anvil stone and striking it hard enough with a hammer stone without causing the nut to fly off of the anvil. The learning process is motivated by many hours of observing a proficient adult performing the act and is shaped during many hours of practice of the individual steps involved (Inoue-Nakamura and Matsuzawa, 1997; Matsuzawa, 1994; Hayashi et al., 2005). The final functional sequence eventually develops, and the skill can be usefully employed to extract the rich and nutritious meat inside the nuts.

For a tool to be functional, the user must learn about its properties and about how the user can manipulate the tool to achieve a goal. This learning might involve only a superficial understanding of the tool. For example, the animal might learn how to use the tool through procedural or habit learning without representing the underlying causal structure of the tool. The tool user may, however, acquire a deeper understanding of how the tool works. They may represent both the tool's physical properties, the rules by which those properties can be put to use (i.e., functional properties and affordances), the goal motivating the use of the tool, and the interrelationships among these domains of knowledge. This is a more complex set of relationships than what the rat was faced with in the experiment by Blaisdell et al. (2006). Perhaps this is why it has been difficult to empirically demonstrate that a tool-using animal understands the causal properties of tools and their effects.

The trap-tube task was first developed to study causal reasoning processes involved in tool use in capuchin monkeys and chimpanzees and has become a standard test for assessing causal understanding involved in tool use in nonhuman animals (Limongelli et al., 1995; Povinelli, 2000; Visalberghi and Limongelli, 1994). The trap-tube task involves the placement of a piece of food inside of a clear tube with two open ends (Fig. 19). The subject is provided with a stick which can be used to retrieve the food by pushing



**Figure 18** Chimpanzee using hammer and anvil stones as tools to crack open palm nuts. With permission of Tetsuro Matsuzawa.

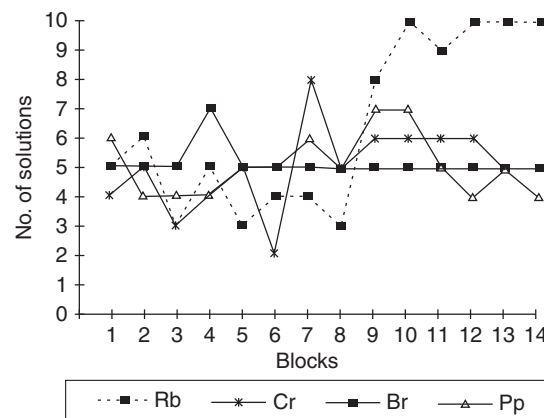


**Figure 19** Trap-tube task for capuchin monkeys (Visalberghi and Limongelli, 1994). See text. Reprinted from Visalberghi, E., Limongelli, L., 1994. Lack of comprehension of cause-effect relations in tool-using capuchin monkeys (*Cebus apella*). J. Comp. Psychol. 108, 15–22, with permission from E. Visalberghi.

the food out of one end of the tube. On one side of the tube there is a hole which can trap the food if it is pushed across it. If the subject understands the nature of the trap, it should push the food out of the opposite end from the side with the trap.

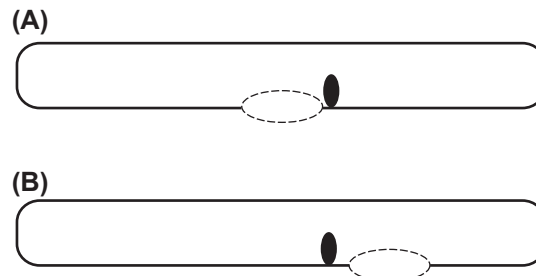
In an initial study, three of four capuchins failed to learn to push the food out of the side of the tube away from the trap (see Fig. 20), and even the successful subject only learned after about 80 trials.

Further tests revealed, however, that the successful subject did not understand how the action of the stick affected the displacement of the reward. Rather, it appeared to solve the task using a simple distance-based associative rule of placing the stick into the side of the tube furthest from the reward. A new trap tube was constructed with the trap placed closer to one end of the tube (Fig. 21B). To prevent food from falling into the trap in Trap Tube B, the subject had to insert the stick into the side of the tube closest to the food rather than the side of the trap furthest from the food as was the correct solution for Trap Tube A. The successful subject, however, continued to insert the stick into the side furthest from the reward in Trap Tube B, hence causing the reward to fall into the trap. This clearly indicates that the successful monkey did not understand the relationship between the trap and the direction the food had to be moved to be successfully retrieved.



**Figure 20** Number of solutions in each 10-trial block of Experiment 1 of Visalberghi and Limongelli (1994). See text. Reprinted from Visalberghi, E., Limongelli, L., 1994. Lack of comprehension of cause-effect relations in tool-using capuchin monkeys (*Cebus apella*). J. Comp. Psychol. 108, 15–22, with permission from E. Visalberghi.





**Figure 21** Trap-tube task for capuchin monkeys (Visalberghi and Limongelli, 1994) and chimpanzees (Limongelli et al., 1995). Trap Tube A is the training configuration with the trap located in the center of the tube and the food placed to one side of the trap. Trap Tube B is the transfer configuration with the trap located closer to one side of the tube and food placed in the same location within the trap that it appeared in the training tube. While Trap Tube A required placing the stick in the side of the tube furthest from the reward to prevent the food from becoming trapped, Trap Tube B required changing this distance-based solution to one of inserting the stick into the side of the tube closest to the food. Reprinted from Visalberghi, E., Limongelli, L., 1994. Lack of comprehension of cause-effect relations in tool-using capuchin monkeys (*Cebus apella*). *J. Comp. Psychol.* 108, 15–22, with permission from E. Visalberghi.

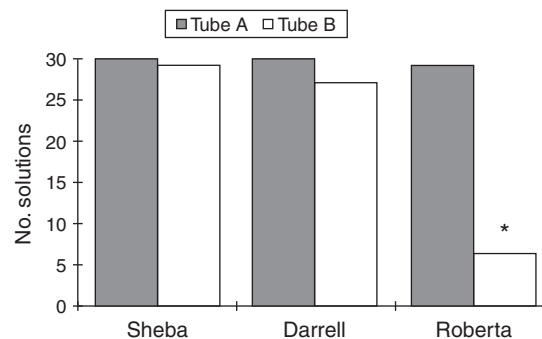
Five chimpanzees were also tested on the trap tube problem, and like the capuchins, only some of the chimpanzees learned to solve the task successfully (Limongelli et al., 1995). Unlike the capuchins, however, the two chimps that learned to solve the task did not appear to be using a simple distance-based strategy. When presented with Trap Tube B (Fig. 21B), the chimps that were successful with Tube A correctly inserted the stick into the side of the trap closest to the food, thereby allowing them to push the food out of the other end of the tube (Fig. 22).

Although the interpretation of how the chimpanzees solved the trap tube problem has been challenged with scrutiny (Povinelli, 2000; Tomasello and Call, 1997), more recent demonstrations that avoid some of the methodological problems of the earlier studies tend to support the existence of some form of appreciation of the causal nature of the task, at least by small samples of apes (Martin-Ordas et al., 2008; Mulcahy and Call, 2006).

#### 1.06.8.2 Tool Use in Corvids

Despite the extensive research on primates engaged in the trap-tube task, some of the strongest evidence that animals can represent the causal structure underlying the tools they use comes from rooks (*Corvus frugilegus*), a member of the Corvid family that includes crows, ravens, and jays (Fig. 23). Although rooks are not known to use tools habitually in the wild, they will readily do so in a laboratory setting. In one study, rooks were presented with a modified trap-tube task (Seed et al., 2006; see also Taylor et al., 2009 for a similar study in the New Caledonian Crow, a tool-using species of corvid).

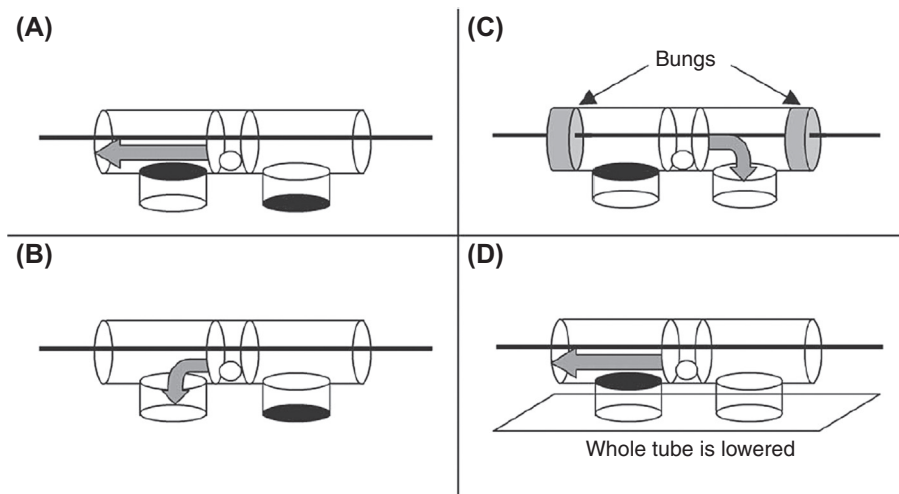
Rooks were presented with the task shown in Fig. 24, Panel A, while others were presented with the task shown in Panel B. Each task involved a tube with a small piece of food located in its center. A stick was already inserted into the tube at the start of the trial, and the food was enclosed by two plastic discs. Pushing the stick forward or pulling it backward would cause the food to move in the direction of the push or pull. A “trap” was located below the floor of both sides of the tube. One of the traps was effective in that if the food was dragged across the top, it would drop inside and become trapped—preventing the subject from retrieving the food. The other trap was ineffective and could not trap the food. For the tube in Panel A, the trap on the right side was functional and could



**Figure 22** Performance of chimpanzees (Sheba and Darrell) and a capuchin monkey (Roberta) on Trap Tube B after having successfully learning to retrieve food from Trap Tube A. Both chimpanzees correctly inserted the stick into the side of the tube closest to the reward, allowing the successful retrieval of the food. The capuchin did not consistently insert the stick into the correct side of Tube B, and hence the food was trapped on most trials. Reprinted from Limongelli, L., Boysen, S.T., Visalberghi, E., 1995. Comprehension of cause-effect relations in tool-using task by chimpanzees (*Pan troglodytes*). *J. Comp. Psychol.* 109, 18–26, with permission from E. Visalberghi.



**Figure 23** A rook. Photo courtesy of Raven.

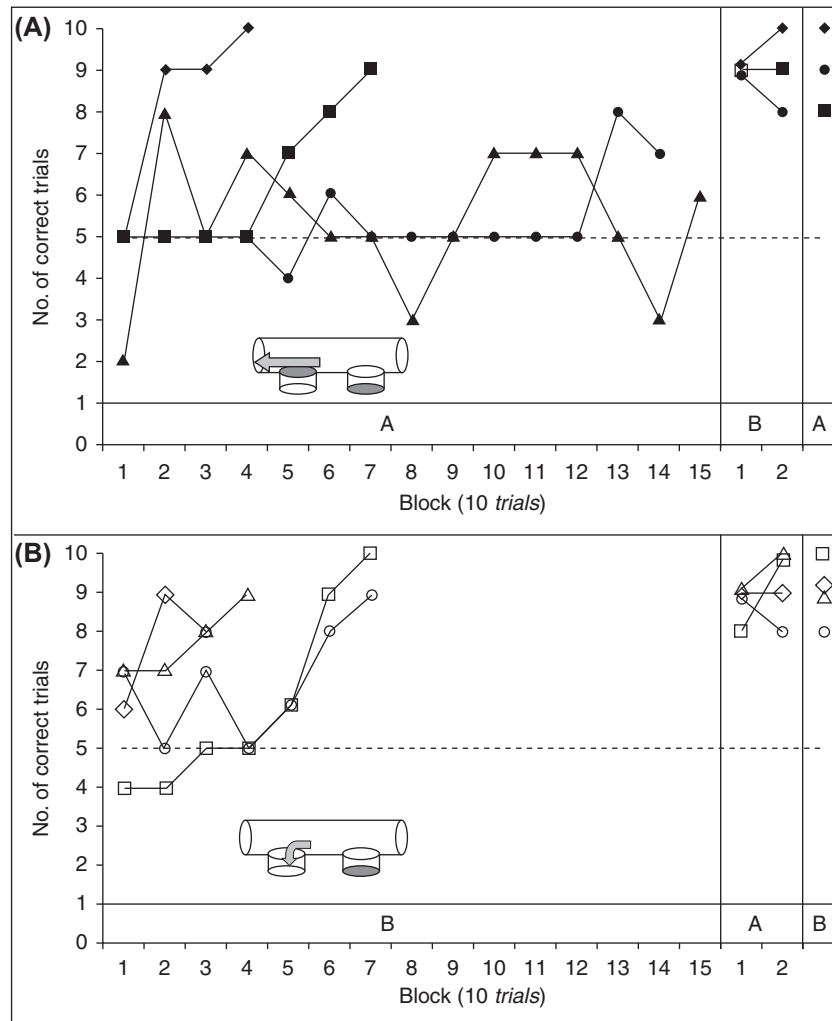


**Figure 24** Trap tube problems used by Seed et al. (2006). See text. Reprinted from Seed, A.M., Tebbich, S., Emery, N.J., Clayton, N.S., 2006. Investigating physical cognition in rooks, *Corvus frugilegus*. *Curr. Biol.* 16, 697–701, © 2006, with permission from Elsevier.

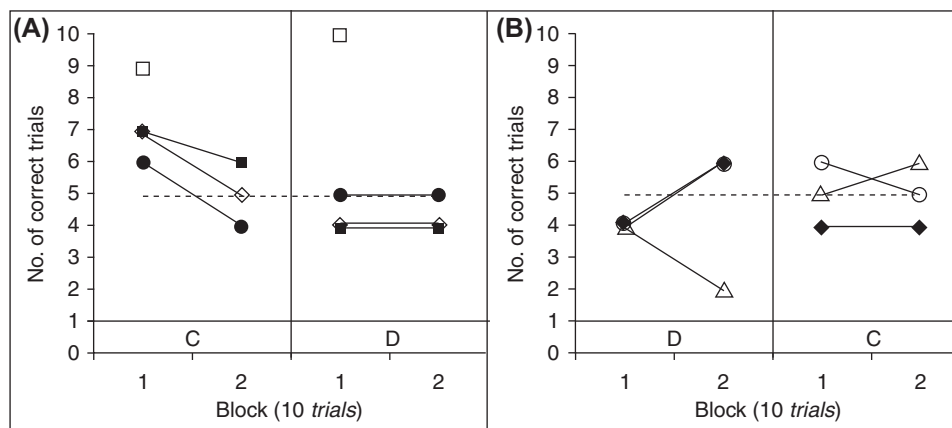
trap the food, while the tube on the left was ineffective. Hence, to solve the tube-trap task shown in Panel A, the subject should move the food over the nonfunctional trap on the left, thereby allowing the food to be retrieved. (Note, the side of the functional trap was left-right counterbalanced across trials to prevent the subject from learning a simple side bias to solve the task.) The tube shown in Panel B contained a similar functional trap to that shown in Panel A, but the nonfunctional trap was different. Rather than pulling the food entirely across the surface of the ineffective trap, the trap was completely open so that food would fall through and out of the tube where the subject could retrieve it.

Seven out of eight rooks learned to solve the task they were given (Fig. 25). Task B was learned more quickly than Task A. Furthermore, transfer was almost perfect to the other task (i.e., from A to B or from B to A). This excellent performance suggests that the rooks acquired a deeper causal understanding of the task. There is, however, a simpler alternative explanation for the superb transfer between tasks. Because both tubes have the same functional trap, it is possible that the birds had simply learned to avoid moving the food in the direction of the functional trap. To test this alternative hypothesis, the birds were presented with the trap tubes shown in Panels C and D of Fig. 24. These tubes lacked the functional trap present in both Tubes A and B, and both tubes contained the nonfunctional traps from Tubes A and B. One of the originally nonfunctional traps from the initial tasks was now made functional in the new tasks. For Tube C, the nonfunctional trap from Tube A was made functional by placing bungs at the ends of the tube. Thus, the food could only be retrieved by dragging it across the trap with the opening in the bottom. For Tube D, the nonfunctional trap from Tube B was made functional by lowering the entire tube onto the floor. Thus, food dragged across the trap from Tube B would trap the food inside, while dragging the food across the trap from Tube A would allow the bird to retrieve the food.

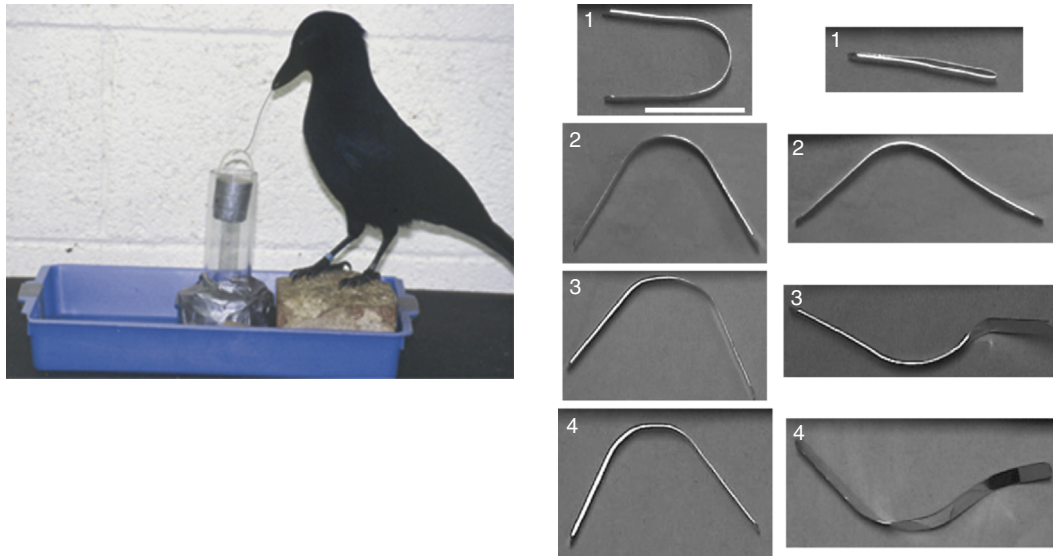
Fig. 26 shows the performance on Tubes C and D. Only one bird (open squares) consistently performed well above chance on these tubes, suggesting that this bird had acquired a deep understanding of the causal structure of the trap tubes. None of the other birds were consistently above chance and thus likely had learned to solve Tubes A and B by avoiding the functional trap. Thus, while a majority of the birds appeared not to understand the causal properties of the tasks, one bird demonstrated knowledge of the causal



**Figure 25** Results of Experiment 1 from Seed et al. (2006). Panel (A) shows those birds that received Tube A in Experiment 1 and Tube B in Experiment 2. Panel (B) shows those that received Tube B first. The horizontal dotted line in each graph shows chance performance (5 out of 10 trials correct). The vertical lines separate results from different tubes, and the boxes above the x axis show which of the tubes were used. Reprinted from Seed, A.M., Tebbich, S., Emery, N.J., Clayton, N.S., 2006. Investigating physical cognition in rooks, *Corvus frugilegus*. Curr. Biol. 16, 697–701, © 2006, with permission from Elsevier.



**Figure 26** Results from Experiment 3 of Seed et al. (2006). Panel (A) shows those birds that received Tube C and then Tube D. Panel (B) shows those that received Tube D first. The horizontal dotted line in each graph shows chance performance (5 out of 10 trials correct). The vertical lines separate results from different tubes, and the boxes above the x axis show which of the tubes were used. Reprinted from Seed, A.M., Tebbich, S., Emery, N.J., Clayton, N.S., 2006. Investigating physical cognition in rooks, *Corvus frugilegus*. Curr. Biol. 16, 697–701, © 2006, with permission from Elsevier.



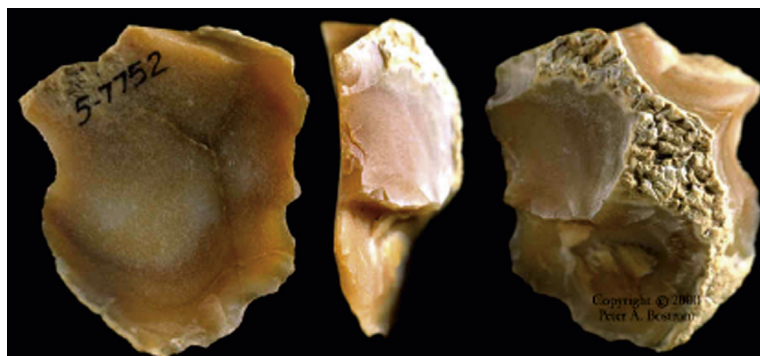
**Figure 27** Left panel: Betty, a New Caledonian crow, using a modified hook to retrieve a food reward. Right panel: Examples of hook tools modified by Betty in an experiment involving the retrieval of food. Photo courtesy of Alex Weir.

properties of objects, such as that objects fall toward the earth when unsupported, that objects cannot move through physical barriers. The generalization of the knowledge about the properties of objects, and of the functional properties of the trap tube itself (i.e., how the food in the tube can be moved by moving the stick), allowed this bird to solve Tasks C and D.

The New Caledonian crow (*Corvus moneduloides*), another species of Corvid, has shown the ability to develop novel techniques to bend aluminum strips to use them as hook tools to retrieve food (Fig. 27). The rapidity with which the crow achieved success shows that the behavior is not only goal directed, but also conveys a deeper understanding of some of the properties of the tools and the materials and a representation of the kinds of solutions that are likely to work (Kenward et al., 2005; Weir et al., 2002). In fact, simple generalization of learned rules could not explain the excellent performance because in most cases the solution required manipulating the tool in a way that was inconsistent or that conflicted with prior successful solutions. For example, one crow named Betty was able to correctly anticipate on four out of five trials whether bending or unbending pieces of novel materials were required to retrieve food (Weir and Kacelnik, 2006). Even this adept crow, however, often probed the recess containing the out-of-reach food with unmodified tools before modifying them or attempted to use the unmodified end of a modified tool. Though suggestive, the details of Betty's performance prevent us from determining whether she understood the physical causality underlying these tools.

### 1.06.8.3 Tool Use by Humankind

Although some animal species, in particular among the corvids and primates, have shown remarkable tool use abilities in the laboratory and the wild, these abilities pale in comparison to that shown by humankind. The earliest evidence for tool use among the hominids dates back by at least 2 million years where modified stone tools consisting of struck flakes of volcanic rock have been found among the remains of our ancestor *Homo habilis* (Fig. 28; Klein, 1989).



**Figure 28** Oldowan flake tools. Olduvai Gorge, Tanzania, Africa. University of California, Berkeley, Department of Anthropology Collection. <http://lithiccastinglab.com/gallery-pages/oldowanflake-triple-large.htm>.

There is more extensive evidence of stone tool use by *Homo erectus* throughout Africa, Asia, and Europe dating from about 1.8 to 0.5 million years ago. By 40,000 years ago, when anatomically modern *Homo sapiens*—historically called Cro-Magnon but more appropriately referred to as *Homo sapiens sapiens*—first arrived at Europe, tool use had become quite sophisticated compared to earlier species of *Homo*. Furthermore, approximately 80,000 to 60,000 years ago, the evolution of tools and other technologies advanced at a very rapid pace and showed major geographical diversification (Brown et al., 2012; Mellars, 2008). Archeological sites throughout Europe show evidence of painting, engraving, sculpture, body ornamentation, and even music. There is even indirect evidence for the weaving of wool into cloth, such as is used for clothing. Remains of dwellings, sculpture, and weapons are plentiful at many archeological sites, but perhaps Cro-Magnons are most famous for their cave paintings.

Humankind's extensive, habitual tool use, along with the evolution of language and extended brain evolution and development, contributes to our unique ability to adapt ourselves to life in almost every niche of the globe, to exploit a wide range of natural resources from engineered crops to nuclear energy, and even allows for our occasional forays into space. Indeed, the innate drive to play during childhood and even into adulthood likely plays a pivotal role in shaping our intelligence by exposing the individual to the S–R–O contingencies of our world, driving neural plasticity and the formation of deeper and more abstract representations of our world (Blaisdell, 2015).

## Acknowledgments

Support for the production of this chapter was provided by NIMH Grant MH 070633 (A.P. Blaisdell). I would like to thank Anthony Dickinson and Michael Waldmann for illuminating discussions that played a seminal role in shaping the story told herein. Thanks are also due to the members of CRITAKAL (Causal Reasoning In The Animal Kingdom Assessment Lab) who further shaped the scope of my thesis. Members include Keith Holyoak, Patricia Cheng, Michael Waldmann, Mimi Liljeholm, Miriam Bassok, Kenneth Leising, and W. David Stahlman. Requests for reprints should be addressed to Aaron P. Blaisdell, UCLA Department of Psychology, 1285 Franz Hall, Box 951563, Los Angeles, CA 90095-1563, USA; email: [Blaisdell@psych.ucla.edu](mailto:Blaisdell@psych.ucla.edu).

**See also:** 1.22 Behavioral Analysis of Learning and Memory in Honeybees. 1.28 Operant Behavior in Model Systems.

## References

- Adams, C.D., Dickinson, A., 1981. Instrumental responding following reinforcer devaluation. *Q. J. Exp. Psychol. Comp. Physiol. Psychol.* 33, 109–121.
- Adams, C.D., 1980. Postconditioning devaluation of an instrumental reinforcer has no effect on extinction performance. *Q. J. Exp. Psychol.* 32, 447–458.
- Adams, C.D., 1982. Variations in the sensitivity of instrumental responding to reinforcer devaluation. *Q. J. Exp. Psychol.* 34B, 77–98.
- Antonitis, J.J., 1951. Response variability in the white rat during conditioning, extinction, and reconditioning. *J. Exp. Psychol.* 42 (4), 273.
- Balleine, B.W., Dickinson, A., 1998. Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology* 37, 407–419.
- Balleine, B.W., Dickinson, A., 1998. The role of incentive learning in instrumental outcome revaluation by sensory-specific satiety. *Anim. Learn. Behav.* 26, 46–59.
- Balleine, B.W., O'Doherty, J.P., 2010. Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology* 35 (1), 48–69.
- Balleine, B., 1992. Instrumental performance following a shift in primary motivation depends on incentive learning. *J. Exp. Psychol. Anim. Behav. Process.* 18, 236–250.
- Balsam, P.D., Deich, J.D., Ohyama, T., Stokes, P.D., 1998. Origins of new behavior. In: O'Donohue, W.T. (Ed.), *Learning and Behavior Therapy*. Allyn & Bacon, Needham Heights, MA, USA, pp. 403–420 xxiii, 568 pp.
- Bays, P.M., Flanagan, J.R., Wolpert, D.M., 2006. Attenuation of self-generated tactile sensations is predictive, not postdictive. *PLoS Biol.* 4, e28.
- Blaisdell, A.P., Waldmann, M.R., 2012. Rational rats: causal inference and representation. In: Wasserman, E.A., Zentall, T.R. (Eds.), *Handbook of Comparative Cognition*. Oxford University Press, Oxford, pp. 175–198.
- Blaisdell, A.P., Sawa, K., Leising, K.J., Waldmann, M.R., 2006. Causal reasoning in rats. *Science* 111 (5763), 1020–1022.
- Blaisdell, A.P., Leising, K.J., Stahlman, W.D., Waldmann, M.S., 2009. Rats distinguish between absence of events and lack of information in sensory preconditioning. *Int. J. Comp. Psychol.* 22, 1–18.
- Blaisdell, A.P., Stolyarova, A., Stahlman, W.D., 2016. The law of expect or a modified law of effect? Outcome expectation and variation in learned behavior. *Conductual Int. J. Interbehaviorism Behav. Anal.* 4 (2), 61–90.
- Blaisdell, A.P., 2003. The S–R stream: where's the filter. *Integr. Physiol. Behav. Sci.* 38, 147–165. PMID:14527183.
- Blaisdell, A.P., 2009. The role of associative processes in spatial, temporal, and causal cognition. In: Watanabe, S., Blaisdell, A.P., Huber, L., Young, A. (Eds.), *Rational Animals, Irrational Humans*. Keio University, Tokyo, Japan, pp. 153–172.
- Blaisdell, A.P., 2015. Play as the foundation of human intelligence: the illuminating role of human brain evolution and development and implications for education and child development. *J. Evol. Health* 1 (1). Article 9. <http://dx.doi.org/10.15310/2334-3591.1016>.
- Bolles, R.C., 1972. Reinforcement, expectancy, and learning. *Psychol. Rev.* 79, 394–409.
- Brogden, W.J., 1939. Sensory pre-conditioning. *J. Exp. Psychol.* 25 (4), 323.
- Brown, K.S., Marean, C.W., Jacobs, Z., Schoville, B.J., Oestmo, S., Fisher, E.C., Bernatchez, J., Karkanas, P., Matthews, T., 2012. An early and enduring advanced technology originating 71,000 years ago in South Africa. *Nature* 491, 590–594.
- Carr, H., Watson, J.B., 1908. Orientation in the white rat. *J. Comp. Neurol. Psychol.* 18 (1), 27–44.
- Chen, X.Y., Wolpaw, J.R., 1995. Operant conditioning of H-reflex in freely moving rats. *J. Neurophysiol.* 73, 411–415.
- Cheng, P.W., 1997. From covariation to causation: a causal power theory. *Psychol. Rev.* 104 (2), 367–405.
- Colwill, R., Rescorla, R.A., 1988. Associations between the discriminative stimulus and the reinforcer in instrumental learning. *J. Exp. Psychol. Anim. Behav. Process.* 14, 155–164.
- Colwill, R.M., Rescorla, R.A., 1985. Postconditioning devaluation of a reinforcer affects instrumental responding. *J. Exp. Psychol. Anim. Behav. Process.* 11, 120–132.
- Colwill, R.M., Rescorla, R.A., 1986. Associative structures in instrumental learning. In: Bower, G.H. (Ed.), *The Psychology of Learning and Motivation*, vol. 20. Academic Press, San Diego, pp. 55–104.
- DeCasper, A.J., Fifer, W.P., 1980. Of human bonding: newborns prefer their mothers' voices. *Science* 208 (4448), 1174–1176.



- Dennett, D.C., 1989. *The Intentional Stance*. Mass.: MIT Press/A Bradford Book, Cambridge.
- Dennett, D.C., 1995. *Darwin's Dangerous Idea: Evolution and the Meanings of Life*. Simon & Schuster, New York.
- Dickinson, A., Balleine, B., 1994. Motivational control of goal-directed action. *Anim. Learn. Behav.* 22, 1–18.
- Dickinson, A., Charnock, D.J., 1985. Contingency effects with maintained instrumental reinforcement. *Q. J. Exp. Psychol. Comp. Physiol. Psychol.* 37, 397–416.
- Dickinson, A., Mulatero, C.W., 1989. Reinforcer specificity of the suppression of instrumental performance on a non-contingent schedule. *Behav. Process.* 19 (1–3), 167–180.
- Dickinson, A., 1980. *Contemporary Animal Learning Theory*. Cambridge University Press, Cambridge.
- Dickinson, A., 1988. Intentionality in animal conditioning. In: Weiskrantz, L. (Ed.), *Thought Without Language*. Clarendon Press/Oxford University Press, New York, NY, USA, pp. 305–325.
- Eckerman, D.A., Lanson, R.N., 1969. Variability of response location for pigeons responding under continuous reinforcement, intermittent reinforcement, and extinction. *J. Exp. Anal. Behav.* 12 (1), 73–80.
- Edwards, D.H., Heitler, W.J., Krasne, F.B., 1999. Fifty years of a command neuron: the neurobiology of escape behavior in the crayfish. *Trends Neurosci.* 22, 153–161.
- Epstein, R., 2014. On the orderliness of behavioral variability: insights from generativity theory. *J. Contextual Behav. Sci.* 3, 279–290.
- Fast, C.D., Blaisdell, A.P., 2011. Rats are sensitive to ambiguity. *Psychon. Bull. Rev.* 18, 1230–1237. PMID:21968926.
- Fast, C.D., Biedermann, T., Blaisdell, A.P., 2016. Imagine that! Cue-evoked representations guide rat behavior during ambiguous situations. *J. Exp. Psychol. Anim. Learn. Cogn.* 42, 200–211.
- Fast, C.D., Flesher, M.M., Nocera, N.A., Fanselow, M.S., Blaisdell, A.P., 2016. Learning history and cholinergic modulation in the dorsal hippocampus are necessary for rats to infer the status of a hidden event. *Hippocampus*. <http://dx.doi.org/10.1002/hipo.22564>.
- Flaherty, C.F., Grisson, P.S., 1988. From contrast to reinforcement: role of response contingency in anticipatory contrast. *J. Exp. Psychol. Anim. Behav. Process.* 14 (2), 165.
- Flaherty, C.F., 1996. *Incentive Relativity*. Cambridge University Press, New York, NY.
- Frick, F.C., Miller, G.A., 1951. A statistical description of operant conditioning. *Am. J. Psychol.* 64 (1), 20–36.
- Gallistel, C.R., Mark, T.A., King, A.P., Latham, P.E., 2001. The rat approximates an ideal detector of changes in rates of reward: implications for the law of effect. *J. Exp. Psychol. Anim. Behav. Process.* 27 (4), 354.
- Gharib, A., Gade, C., Roberts, S., 2004. Control of variation by reward probability. *J. Exp. Psychol. Anim. Behav. Process.* 30, 271–282.
- Gopnik, A., Schulz, L., 2004. Mechanisms of theory formation in young children. *Trends Cogn. Sci.* 8 (8), 371–377.
- Gopnik, A., Glymour, C., Sobel, D.M., Schulz, L.E., Kushnir, T., Danks, D., 2004. A theory of causal learning in children: causal maps and bayes nets. *Psychol. Rev.* 111, 3–32.
- Haggard, P., Clark, S., 2003. Intentional action: conscious experience and neural prediction. *Conscious Cogn.* 12, 695–707.
- Haggard, P., Clark, S., Kalogeras, J., 2002. Voluntary action and conscious awareness. *Nat. Neurosci.* 5, 382–385.
- Haggard, P., 2005. Conscious intention and motor cognition. *Trends Cogn. Sci.* 9, 290–295.
- Hammond, L.J., 1980. The effect of contingency upon the appetitive conditioning of free-operant behavior. *J. Exp. Anal. Behav.* 34, 297–304.
- Hayashi, M., Mizuno, Y., Matsuzawa, T., 2005. How does stone-tool use emerge? Introduction of stones and nuts to naive chimpanzees in captivity. *Primates* 46, 91–102.
- Herrick, R.M., Bromberger, R.A., 1965. Lever displacement under a variable ratio schedule and subsequent extinction. *J. Comp. Physiol. Psychol.* 59 (3), 392.
- Heyes, C., Dickinson, A., 1990. The intentionality of animal action. *Mind Lang.* 5, 87–104.
- Holland, P.C., 1979. Differential effects of omission contingencies on various components of pavlovian appetitive conditioned responding in rats. *J. Exp. Psychol. Anim. Behav. Process.* 5, 178–193.
- Inoue-Nakamura, N., Matsuzawa, T., 1997. Development of stone tool use by wild chimpanzees (*Pan troglodytes*). *J. Comp. Psychol.* 111, 159–173.
- Johanson, I.B., Hall, W.G., 1979. Appetitive learning in 1-day-old rat pups. *Science* 205 (4404), 419–421.
- Kenward, B., Weir, A.A., Rutz, C., Kacelnik, A., 2005. Behavioural ecology: tool manufacture by naive juvenile crows. *Nature* 433 (7022), 121.
- Kipling, R., 1912. *Just so Stories*. Doubleday, Garden City, NY.
- Klein, R.G., 1989. *The Human Career: Human Biological and Cultural Origins*. The University of Chicago Press, Chicago.
- Konorski, J., Miller, S., 1936. Conditioned reflexes of the motor analyser. In: Trudi Fiziologicheskikh Laboratoriyi Akademika J. P. Pavlova, vol. 6, pp. 119–278 (In Russian with English Summary, pp. 285–288).
- Kruse, J.M., Overmier, J.B., Konz, W.A., Rokke, E., 1983. Pavlovian conditioned stimulus effects upon instrumental choice behavior are reinforcer specific. *Learn. Motiv.* 14 (2), 165–181.
- Leising, K.J., Wong, J., Waldmann, M.S., Blaisdell, A.P., 2008. The special status of actions in causal reasoning in rats. *J. Exp. Psychol. Gen.* 137, 514–527. PMID:18729713.
- Leslie, A.M., 1995. A theory of agency. In: Sperber, D., Premack, D., Premack, A.J. (Eds.), *Causal Cognition: A Multidisciplinary Debate*. Clarendon Press/Oxford University Press, New York, NY, pp. 121–149.
- Libet, B., 1985. Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behav. Brain Sci.* 8, 529–566.
- Limongelli, L., Boysen, S.T., Visalberghi, E., 1995. Comprehension of cause-effect relations in a tool-using task by chimpanzees (*Pan troglodytes*). *J. Comp. Psychol.* 109, 18–26.
- Mackintosh, N.J., Dickinson, A., 1979. Instrumental (Type II) conditioning. In: Dickinson, A., Boakes, R.A. (Eds.), *Mechanisms of Learning and Motivation*. Erlbaum, Hillsdale, NJ, pp. 143–167.
- Marr, D., 1982. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. W.H. Freeman, San Francisco.
- Martin-Ordas, G., Call, J., Colmenares, F., 2008. Tubes, tables and traps: great apes solve two functionally equivalent trap tasks but show no evidence of transfer across tasks. *Anim. Cogn.* 11 (3), 423–430.
- Matsuzawa, T., 1994. Field experiments on use of stone tools by chimpanzees in the wild. In: Wrangham, R.W., McGrew, W.C., de Waal, F.B.M., Heltne, P.G. (Eds.), *Chimpanzee Cultures*. Harvard University Press, Cambridge, MA, pp. 351–370.
- Mellars, P., 2008. Why did modern human populations disperse from Africa ca. 60,000 years ago? A new model. *Proc. Natl. Acad. Sci. U.S.A.* 103, 9381–9386.
- Millenson, J.R., Hurwitz, H.M.B., 1961. Some temporal and sequential properties of behavior during conditioning and extinction. *J. Exp. Anal. Behav.* 4, 97–106.
- Miller, S., Konorski, J., 1928. Sur une forme particuliere des reflexes conditionnels. *Compte Rendu Hebdomadaire des Seances Mtrmoires de la Soc. de Biol.* 99, 1151–1157.
- Mulcahy, N.J., Call, J., 2006. How great apes perform on a modified trap-tube task. *Anim. Cogn.* 9, 193–199.
- Nadler, N., Delgado, M.R., Delamater, A.R., 2011. Pavlovian to instrumental transfer of control in a human learning task. *Emotion* 11 (5), 1112–1123.
- Neuringer, A., Kornell, N., Olufs, M., 2001. Stability and variability in extinction. *J. Exp. Psychol. Anim. Behav. Process.* 27, 79–94.
- Notterman, J.M., 1959. Force emission during bar pressing. *J. Exp. Psychol.* 58, 341–347.
- Pavlov, I.P., 1927. *Conditioned Reflexes*. Oxford University Press, Oxford, UK.
- Pearl, J., 2000. *Causality: Models, Reasoning, and Inference*. Cambridge University Press, Cambridge, UK.
- Povinelli, J., 2000. *Folk Physics for Apes: The Chimpanzee's Theory of How the World Works*. Oxford University Press, Oxford/New York.
- Premack, D., 1990. The infant's theory of self-propelled objects. *Cognition* 36, 1–16.
- Racey, D., Young, M.E., Garlick, D., Pham, J.N.-M., Blaisdell, A.P., 2011. Pigeon and human performance in a multi-armed bandit task in response to changes in variable interval schedules. *Learn. Behav.* 39, 245–258. PMID:21380732.
- Rescorla, R.A., Solomon, R.L., 1968. Two-process learning theory: relationships between Pavlovian conditioning and instrumental learning. *Psychol. Rev.* 74, 151–182.
- Schmahmann, J.D., 1991. An emerging concept: the cerebellar contribution to higher function. *Arch. Neurol.* 48 (11), 1178–1187.
- Seed, A.M., Tebbich, S., Emery, N.J., Clayton, N.S., 2006. Investigating physical cognition in rooks, *Corvus frugilegus*. *Curr. Biol.* 16, 697–701.
- Sirigu, A., Daprati, E., Ciancia, S., Giraux, P., Nighoghossian, N., Posada, A., et al., 2004. Altered awareness of voluntary action after damage to the parietal cortex. *Nat. Neurosci.* 7, 80–84.



- Skinner, B.F., 1966. The phylogeny and ontogeny of behavior. *Science* 153, 1205–1213.
- Sloman, S., 2005. *Causal Models: How People Think about the World and Its Alternatives*. Oxford University Press, Oxford.
- Spirtes, P., Glymour, C., Scheines, R., 1993. *Causation, Prediction, and Search*. Springer-Verlag, New York.
- Stahlman, W.D., Blaisdell, A.P., 2011. Reward probability and the variability of foraging behavior in rats. *Int. J. Comp. Psychol.* 24, 168–176.
- Stahlman, W.D., Blaisdell, A.P., 2011. The modulation of response variation by the probability, magnitude, and delay of reinforcement. *Learn. Motiv.* 42, 221–236.
- Stahlman, W.D., Roberts, S., Blaisdell, A.P., 2010. Effect of reward probability on spatial and temporal variation. *J. Exp. Psychol. Anim. Behav. Process.* 36, 77–91.
- Stahlman, W.D., Young, M.E., Blaisdell, A.P., 2010. Response variability in pigeons in a Pavlovian task. *Learn. Behav.* 38, 111–118.
- Stebbins, W.C., Lanson, R.N., 1962. Response latency as a function of reinforcement schedule. *J. Exp. Anal. Behav.* 5, 299–304.
- Steyvers, M., Tenenbaum, J.B., Wagenmakers, E., Blum, B., 2003. Inferring causal networks from observations and interventions. Special Issue: 2002 Rumelhart Prize Special Issue Honoring Richard Shiffrin *Cogn. Sci.* 27, 453–489.
- Sutherland, R.J., Rudy, J.W., 1989. Configural association theory: the role of the hippocampal formation in learning, memory, and amnesia. *Psychobiology* 17 (2), 129–144.
- Taylor, A.H., Hunt, G.R., Medina, F.S., Gray, R.D., 2009. Do New Caledonian crows solve physical problems through causal reasoning? *Proc. R. Soc. Lond. B Biol. Sci.* 276 (1655), 247–254.
- Thorndike, E.L., 1898. *Animal intelligence: an experimental study of the association processes in animals*. Psychol. Rev. Monogr. 2 (Whole No. 8).
- Thorndike, E.L., 1911. *Animal Intelligence: Experimental Studies*. Macmillan, New York.
- Tinklepaugh, O.L., 1928. An experimental study of representative factors in monkeys. *J. Comp. Psychol.* 8, 197–236.
- Tolman, E.C., Gleitman, H., 1949. Studies in learning and motivation: I. equal reinforcements in both end-boxes, followed by shock in one end-box. *J. Exp. Psychol.* 39, 810–819.
- Tolman, E.C., Honzik, C.H., 1930. Introduction and removal of reward, and maze performance in rats. *Univ. Calif. Publ. Psychol.* 4, 257–275.
- Tolman, E.C., 1932. *Purposive Behavior in Animals and Men*. The Century Co, New York.
- Tolman, E.C., 1948. Cognitive maps in rats and men. *Psychol. Rev.* 55, 189–208.
- Tomasello, M., Call, J., 1997. *Primate Cognition*. Oxford University Press, New York.
- Vilis, T., Hore, J., 1980. Central neural mechanisms contributing to cerebellar tremor produced by limb perturbations. *J. Neurophysiol.* 43 (2), 279–291.
- Visalberghi, E., Limongelli, L., 1994. Lack of comprehension of cause-effect relations in tool-using capuchin monkeys (*Cebus apella*). *J. Comp. Psychol.* 108, 15–22.
- Waldmann, M.R., Hagmayer, Y., 2005. Seeing versus doing: two modes of accessing causal knowledge. *J. Exp. Psychol. Learn. Mem. Cogn.* 31, 216–227.
- Waldmann, M.R., Hagmayer, Y., Blaisdell, A.P., 2006. Beyond the information given: causal models in learning and reasoning. *Curr. Dir. Psychol. Sci.* 15, 307–311.
- Waldmann, M.R., Cheng, P.W., Hagmayer, Y., Blaisdell, A.P., 2008. Causal learning in rats and humans: a minimal rational model. In: Chater, N., Oaksford, M. (Eds.), *The Probabilistic Mind: Prospects for Rational Models of Cognition*. Oxford University Press, Oxford, pp. 453–484.
- Waldmann, M.R., Schmid, M., Wong, J., Blaisdell, A.P., 2012. Rats distinguish between absence of events and lack of evidence in contingency learning. *Anim. Cogn.* 15, 979–990.
- Weir, A.A., Kacelnik, A., 2006. A New Caledonian crow (*Corvus moneduloides*) creatively re-designs tools by bending or unbending aluminium strips. *Anim. Cogn.* 9, 317–334.
- Weir, A.A., Chappell, J., Kacelnik, A., 2002. Shaping of hooks in New Caledonian crows. *Science* 297 (5583), 981.
- White, N.M., 1989. Reward or reinforcement: what's the difference? *Neurosci. Biobehav. Rev.* 13, 181–186.
- Wine, J.J., Krasne, F.B., 1972. The organization of escape behavior in the crayfish. *J. Exp. Biol.* 56, 1–18.
- Woodward, J., 2003. *Making Things Happen: A Theory of Causal Explanation*. Oxford University Press, Oxford.
- Yin, H.H., Knowlton, B.J., Balleine, B.W., 2004. Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *Eur. J. Neurosci.* 19, 181–189.