Rational Accounts of Animal Behaviour?
Lessons from C. Lloyd Morgan's Canon

Dominic M. Dwyer & Katy V. Burgess
Cardiff University

One particular concern of the 2010 Winter Conference on Animal Learning and Behaviour was the degree to which the behaviours of human and nonhuman animals might be interpreted as the result of the same cognitive mechanisms. Here, we examine three examples in rats (causal-reasoning, sensitivity to the absence of stimuli, and the relationship between effort and reward) where higher-order mental processes might be invoked as explanations of the observed behaviour. In each case we argue that alternative accounts, based on "lower" mental processes, are also consistent with the observed data. On the basis of the principle of parsimony, enshrined as a grounding assumption of comparative psychology in C. Lloyd Morgan's Canon, the existence of such alternative accounts means that the available evidence does not licence the conclusion that non-human animals display evidence of human-like cognitive processes in these areas.

A central concern for anyone interested in the mental life of nonhuman animals is the degree to which the behaviours of human and nonhuman animals are underpinned by the same cognitive mechanisms. The focus session for the 2010 Winter Conference of Animal Learning and Behavior (WCALB), and this special issue, "The Interface between Learning and Cognition," was directed at investigating this very concern. Within artificial intelligence, neuroscience, and psychology, a dominant assumption has been that nonhuman animals are convenient systems in which to study simple processes (e.g., of learning and memory) untrammelled by the more complex reasoning and rule-based processes possessed by humans. That is, nonhuman animals are basically complex automata - albeit ones that often exhibit a remarkable degree of flexibility. However, this assumption is not universal (especially if the field of comparative psychology is broadly conceived) and intriguing recent research suggests that animals might, in addition to simpler associative processes, also have far richer ways of representing the texture of their environment than has hitherto been supposed (e.g., Blaisdell, Sawa, Leising, & Waldmann, 2006; Foote & Crystal, 2007; Murphy, Mondragon, & Murphy, 2008).

Although initial assumptions regarding the issue of whether nonhuman animals exhibit evidence of complex cognitive mechanisms differ greatly, the field of comparative psychology does share a common intellectual heritage. In particular, Morgan's Canon, "In no case may we interpret an action as the outcome of a higher psychical faculty, if it can be interpreted as the outcome of the exercise of one that stands lower in the psychological scale" (Morgan, 1894, p. 53), represents the fact that due consideration of parsimony is central to the investigation and interpretation of animal behaviour. The idea that animals might be capable of engaging in complex cognition certainly raises the suggestion that

their actions might be the outcome of particularly high psychical facilities. Therefore this paper will consider three examples of behaviour in rats, which were discussed at the 2010 WCALB in terms of complex cognitive mechanisms, to ask whether simpler mechanisms can indeed be ruled out as accounts of these behaviours and thus to examine whether Morgan's Canon has been satisfied.

Causal Reasoning in Rats

One long tradition of research has sought to investigate the capabilities of a variety of species (in particular birds and non-human primates) with respect to causal reasoning via tests of their ability to understand cause-effect relations in tool use (e.g., Fujita, Kuroshima, & Asai, 2003; Martin-Ordas, Call, & Colmenares, 2008; Santos, Pearson, Spaepen, Tsao, & Hauser, 2006; Tebbich & Bshary, 2004; Visalberghi & Limongelli, 1994). Unfortunately this research is currently inconclusive because the results have been somewhat inconsistent. Moreover, in many cases behaviours consistent with true understanding of cause-effect relationships was also consistent with learning by trial and error. The development by Blaisdell et al. (2006) of a novel technique for assessing the capacity for causal reasoning in rats, using a very basic behavioural repertoire, was a major advance from such inconclusive studies based on behaviourally complex tasks. Their task was developed from an analysis of causal model theory (for a review, see Waldmann, Hagmayer, & Blaisdell, 2006) and was effectively a rat-based version of the type of reasoning task used in humans by Waldmann and Hagmayer (2005).

An evocative example will serve to illustrate the nature of the technique developed by Blaisdell et al. (2006). When a single cause (e.g., raindrops) can have two separate effects (wet windows and clothes on a washing line becoming damp), the observation of one effect (wet windows) can lead to the inference that the other effect (damp clothes) has also occurred (this example is taken from Clayton & Dickinson, 2006). Causal model theory suggests that this inference reflects the assumption that the observed effect (wet windows) must be the result of its usual cause (rain), which should also result in its second effect (damp clothes). However, if there is an alternative cause for the first effect (e.g., a water sprinkler near the window, but far from the washing) then the presence of the normal cause (rain) cannot be inferred and thus the other effects of rain are not presumed to occur either. It is this sensitivity to the difference between simply observing an effect and that effect having an alternative cause which Blaisdell et al. (2006) have recently studied in rats (see also, Leising, Wong, Waldmann, & Blaisdell, 2008).

In one experiment, training was given in which a visual stimulus (A: analogous to raindrops) served as the "common cause" for two separate effects: an auditory stimulus, B (wet window), and the delivery of sucrose into a food well (damp clothes). This common cause treatment was arranged by giving rats separate A→B and A→sucrose trials. On further control trials, the rats received another auditory stimulus C (the "direct-cause") that was presented simultaneously with sucrose (i.e., C+sucrose trials). During subsequent testing, the tendency of rats to approach the food well ("nosepoking"; analogous to removing clothes from the line) was examined as a function of whether the presentation of B or C was contingent upon rats pressing a lever (intervene condition) or simply occurred

without apparent cause (observe condition). The rats were less likely to nosepoke if their lever presses resulted in the presentation of B than if B arrived unannounced, but there was no influence of the intervene/observe manipulation on nosepoking during C. This pattern of responding is clearly consistent with causal model theory, and so Blaisdell et al. (2006) concluded that the rats were engaging in causal inference because they could not identify a simpler account of the behaviours observed.

Clearly, the process by which Blaisdell et al. (2006) reached the conclusion that rats are engaging in a process of causal reasoning fits the form of Morgan's Canon: a complex cognitive account was only accepted once simpler accounts were considered and rejected. The question thus becomes whether the rejection of simpler accounts was warranted on the basis of the observed behaviours. We have previously argued that it was not (Dwyer, Starns, & Honey, 2009). This argument was based on a consideration of the ergonomics of the behavioural task presented to the rats: the alternative cause of the auditory stimulus was a lever press, a manipulandum that was separated from the food magazine, and so any tendency to interact with the lever would interfere with any tendency to interact with the food magazine. That is, competition between responses could have contributed to the lower levels of magazine responding during the auditory stimulus when it occurred as a consequence of a rat's own actions. If correct, this very simple account suggests that, *across — but not necessarily within — experimental conditions*, there should be a reciprocal relationship between the tendency to press the lever and the tendency to enter the food magazine during the auditory stimulus. It should be stressed that, in order for response competition to explain differences between conditions in the levels of nosepoking it is response levels across conditions that should be reciprocal: the relationship between responses within an experimental condition might simply reflect individual differences in overall activity or the tendency to interact with environmental stimuli generally. The original data do not speak to the possibility of response competition because lever press responses were not recorded (for all conditions) as a function of whether they occurred during the auditory stimuli. Therefore we (Dwyer, Starns et al., 2009) repeated the procedures used by Blaisdell et al. (2006) while recording lever press responses in a manner that allowed a direct test of response competition. In short, a reciprocal relationship was found across experimental conditions whereby food magazine responding was low when lever pressing was high during the critical auditory stimuli. Unfortunately, we were not able to replicate the theoretically vital interaction from the original study whereby the presentation of the auditory stimulus as a consequence of the rats' own actions only interfered with magazine responding in the common-cause conditions (but see Leising et al., 2008). Thus, while there is direct evidence that lever press and food-magazine responses compete with each other in this type of study (and so the idea that response competition might contribute to the observed behaviour is plausible) we know of no existing experiment where this possibility can be directly tested as an account of the interaction between the effects of the manipulation of training condition (common-cause versus direct cause or causal-chain) and the test condition (observe versus intervene) that is purported to uniquely reflect the operation of causal reasoning.

It should be noted that the existence of competition between lever press and magazine responses does not, in itself, directly invalidate the claim that rats might be capable of causal reasoning. To return to the human example discussed above, the fact that someone could not both turn on a sprinkler at the front of their house and remove the washing at the back at the same time does not mean that they do not understand that there is no need to bring in the washing because it is the sprinkler is wetting the windows and not rain. It is invoking the principle of parsimony instantiated in Morgan's Canon which encourages the acceptance of the simpler account of the rats' behaviour. We would also note that, although we have only considered one potential alternative explanation of the behaviour in the current tasks, other accounts may exist. In particular, new tests of an existing associative model of learning (Schmajuk, Lam, & Gray, 1996) suggest that this model can, without post-hoc modification of its basic structure, be used to provide simulations of the data considered here (Kutlu & Schmajuk, in press). At the time of writing the final report of these new simulations was not publicly available, and so detailed consideration of them is not appropriate here. However, the model under consideration (Schmajuk et al., 1996) instantiates a number of associative mechanisms and its predictions can only be discovered by extensive simulation. That is, while the model is certainly "simple" in that it does not include any contribution of higher cognitive processes, the combination of multiple mechanisms means that it is computationally quite complex. The computational complexity of the Schmajuk et al. (1996) brings into sharp relief the potential tension between different conceptions of parsimony in relation to mental processes and raises the question of whether or not a small number of higher-order mental processes is more complex than a large number of lower-order processes. We will return to this issue later. Regardless, the fact that a purely mechanistic model based on well-established associative principles can mimic data that have been claimed to be uniquely consistent with causal reasoning offers at the very least an existence proof for another account of this data that will require consideration according to Morgan's Canon.

Patterning and Informative Absence

Although somewhat separate from that main focus of the study, the supplementary materials for Blaisdell et al.'s (2006) experiments contained a brief note regarding a serendipitous observation on performance in one of the control conditions. That is, in a condition referred to as "causal-chain," animals were exposed to events such that an auditory cue predicted a light, and that in turn the light predicted food. Such treatment is often referred to as sensory preconditioning (e.g., Rescorla, 1980; Rizley & Rescorla, 1972) and typically results in the observation that animals will make responses appropriate to the anticipation of the reinforcer (e.g., entering the magazine) when presented with the auditory cue, even though it has itself never been directly paired with the reinforcer. In this case, however, animals did not initially make magazine responses when presented with the auditory cue alone, but in subsequent experiments, where the stimulus light was removed from the experimental chamber prior to test, they did. Blaisdell and his colleagues (Blaisdell, Leising, Stahlman, & Waldmann, 2009) suggested that the

difference between the unsuccessful pilot study and the subsequent experiments might be directly related to the removal of the light from the chamber because the rats had acquired a causal chain whereby the auditory stimulus caused the light and the light caused the food and thus the absence of the light would break the causal chain. Blaisdell et al. (2009) provided further evidence consistent with the idea that removing the light prior to test would increase responding to the auditory stimulus in the manner suggested by the previous pilot data. In order to explain why manipulating access to the light would moderate responding to the tone they proposed that the rats were sensitive to the difference between the explicit physical absence of an event and the absence of information about it. Furthermore, they suggested that, in the absence of access to the physical source of a cue, its presence or absence would be considered to be ambiguous and that this ambiguity would influence their expectancies regarding the occurrence of other events that had previously followed the now-absent cue (although they did note that other accounts of the data would be possible). We will return to consider just such an associative account later. Beforehand, we will turn to work regarding patterning procedures that was inspired by the idea that rats might be sensitive to ambiguity.

Fast and Blaisdell (in press) report two studies examining the effect of manipulating the explicit absence of cues on performance following training on patterning discriminations (some of these data were presented by Aaron Blaisdell as part of his keynote address to the WCALB). In Experiment 1, initial training was given in two conditions: either two cues A and B (a solid light on one side of the chamber and a flashing light on the other) were presented such that either the individual cues predicted reinforcement but the combination of both cues did not ("negative pattering": A+, B+, AB-); or the individual cues did not predict reinforcement but the combination did ("positive patterning": A-, B-, AB+). These stimuli were used as discriminative cues indicating when a lever press would be reinforced in an instrumental design. Clear discriminative responding was observed: rats in the positive patterning conditions responded more to the two cues presented together (i.e., AB) than they did to the cues presented alone (A or B), while animals in the negative patterning conditions responded more to the cues presented alone than to the combination. Following the acquisition of discriminative responding, animals' responses to a single cue (A) were examined as a factor of whether the light that represented the other cue (B) was either present (but unilluminated) or absent (in this condition the stimulus light was covered). That is, in the uncovered test conditions, rats had access to both the illuminated cue A and the unilluminated bulb that corresponded to cue B, while in the covered test conditions, rats had the same access to cue A but were unable to see the bulb that corresponded to cue B, and thus could not have ascertained whether it was illuminated or not. Overall, responding to the single cue A was higher in animals given negative patterning than in those given positive pattering and the difference between patterning conditions was reduced when the bulb for cue B was covered. The effect of covering the stimulus bulb was due to a reduction in responding in the negative pattering animals, while those trained on positive pattering were relatively unaffected.

In Experiment 2 all animals were trained on both positive and negative patterning discriminations. One of these was based on the same visual cues as in

Experiment 1 while the other was based on auditory stimuli (the nature of the cues used for each type of discrimination was counterbalanced across animals). Following acquisition, all animals were tested with the visual stimuli as in Experiment 1: that is, they were tested for their response to stimulus A alone (which had been part of either a negative or a positive pattering discrimination) as a factor of whether the stimulus light for B was covered or not. Again, responding to the single cue A was higher in the negative patterning than the positive pattering conditions, and this difference was attenuated when B was covered. Unlike Experiment 1, in Experiment 2, covering B influenced both the positive and negative patterning cases: that is, in the positive pattering case responding to A was higher when B was covered than when it was uncovered, while in the negative pattering case responding to A was lower when B was covered than when it was uncovered. In line with the theoretical discussion of the sensory preconditioning experiments, Fast and Blaisdell (in press) interpreted this pattern of results in light of the idea that the rats were behaving as if they had noticed that the second stimulus light was not available for inspection, concluded that its status was ambiguous, as opposed to simply being not presented, and thus respond in an intermediate fashion that reflected both possible states of the absent cue[1]. Moreover, in order to explain why this modulation of responding was not observed for the positive pattering case in Experiment 1 it was suggested that the concurrent exposure to the negative patterning discrimination might have influenced the manner in which either the discrimination was solved. While the suggestion that rats were sensitive to ambiguity is certainly consistent with the observed data, Morgan's Canon requires that simpler explanations be considered.

This preceding account of the animals' behaviour implies, firstly, that the status of a cue must be represented at least two potential states, on (e.g., A^{on}) and off (e.g., A^{off}), where off corresponds to a light that is present, but unilluminated. At first glance, such A^{off} representations seem to require that rats would be able to explicitly represent the absence of an event. However, it should be remembered that the visual stimuli used by Fast and Blaisdell (in press) consisted of the illumination of a stimulus light and so the absence of the cue actually corresponds to a physically present stimulus — namely the unilluminated bulb. Thus, a consideration of CS^{off} representations does not lead to the absolute requirement that rats are capable of representing cues that are not physically present. That said, the idea that rats might explicitly represent the absence of an event is relatively uncontroversial: it is not uncommon for the absence of an otherwise presented cue to be used as a signal in conditioning experiments (for example Hall & Honey, 1989) used the darkening of a houselight as a cue) and the phenomena of representation mediated conditioning demonstrates that physically absent events can be represented and support learning (Dickinson & Burke, 1996; Dwyer, 2003; Holland, 1990). Thus there is little reason to question the idea that rats might represent either the non-illumination of a physically present stimulus light or even

---

[1]It should be noted that Fast and Blaisdell (in press) are careful to explain that they are not actually attributing awareness of the relevant conditions to the rats, merely that their behaviour served the function of being sensitive to ambiguity. We will return to this point once we have described an associative account of the effects of preventing access to a stimulus.

the absence of an otherwise expected stimulus. However, the preceding account also suggests that preventing access to the source of a cue (i.e., placing an opaque covering over the light bulb) results in the status of the cue being treated as ambiguous. But is some representation of "uncertainty" regarding the status of an absent cue required to explain the behaviour of the rats on this task? That is, if we allow that rats represent both the illumination and non-illumination of the critical stimulus lights, then it is possible to examine the effects of including such CS^{off} representations on associative accounts of the current patterning designs that do not require any consideration of ambiguity.

In order to do this we utilised the ALTSim program (Thorwart, Schultheis, Konig, & Lachnit, 2009) which was developed to provide simulations of common associative models. Stimulus representations were created for A^{on}, A^{off}, B^{on}, and B^{off} and the training procedures modelled such that A trials would involve presenting A^{on} and B^{off}, B trials comprised B^{on} and A^{off}, compound trials comprised A^{on} and B^{on}, while inter-trial intervals comprised A^{off} and B^{off}. Simulation of positive patterning was performed by pairing compound trials (A and B together) with a reinforcer and element (either A or B alone) unreinforced, while negative patterning was instantiated by presenting element trials with the reinforcer and compound trials without reinforcement. In both cases inter-trial intervals (A^{off} and B^{off} together) were presented unreinforced. Training was continued until associative strengths approached asymptote. The test phase was modelled by comparing the associative strength of A^{on} and B^{off} combined (uncovered test) with the associative strength of A^{on} alone (covered test – so neither B^{on} nor B^{off} were presented). Note that simulating the covered test as simply not activating either of the B^{on} or B^{off} representations (because the rats would have access to neither the illuminated nor the unilluminated stimulus light) is not requiring that the status of B is treated as ambiguous, merely that there is no relevant stimulation to excite either the representation of its presence or absence.

Figure 1 (Panels A and B for training, panels C and D for test) shows examples of these simulations according to the configural model proposed by Pearce (1994) and the, elemental, Rescorla-Wagner model (1972). Inspection of Panel A reveals that the addition of CS^{off} representations does not interfere with the ability of Pearce's (1994) model to account for positive and negative patterning. Moreover, Panel B indicates that the addition of such CS^{off} representations does not interfere with the ability of the Rescorla-Wagner model (1972) to account for positive patterning, but nor does it overcome the failure of this model to account for negative patterning. Inspection of Panel C (Pearce's configural model) indicates that activation of the US node by stimulus A was greater in the negative than the positive pattering case and that this difference was reduced when the contribution of Boff representations was removed to simulate the covering of the light for stimulus B. That is exactly the pattern of results as obtained in Experiment 2, and the negative pattering case from Experiment 1, of Fast and Blaisdell (in press). Panel D indicates that the inclusion or exclusion of B^{off} representations has little effect on the predictions of the Rescorla-Wagner model regarding responding to cue A. That is the pattern of results from the positive pattering condition from Experiment 1 of Fast and Blaisdell (in press).
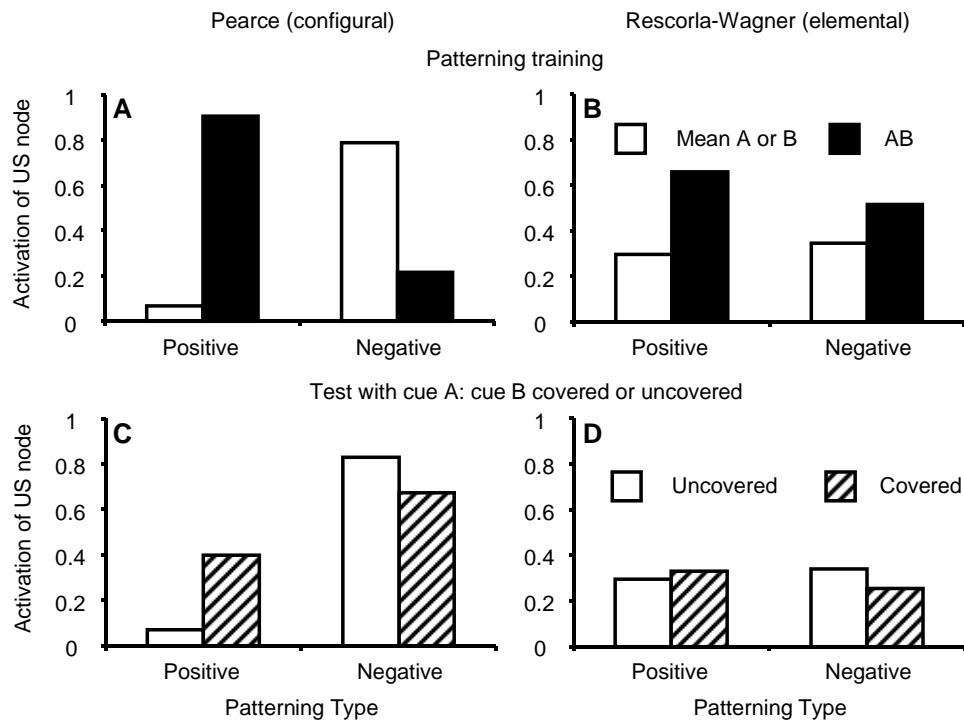
Pearce (configural)   Rescorla-Wagner (elemental)

Patterning training

Test with cue A: cue B covered or uncovered

Figure 1. Shows simulations of Pearce's (1994) configural theory and the (elemental) Rescorla-Wagner model (Rescorla & Wagner, 1972). Panels A (Pearce) and B (Rescorla-Wagner) show the simulation of positive (i.e., A-, B-, AB+) and negative (A+, B+, AB-) pattering. Panels C (Pearce) and D (Rescorla-Wagner) show the simulation of the test phase with cue A as a factor of whether the stimulus light for B was covered or uncovered. As noted in the text, these simulations assume the existence of cues for A^{on}, A^{off}, B^{on}, and B^{off}. For the training simulations, AB reflects the activation of cues A^{on} and B^{on}, while A reflects the activation of cues A^{on} and B^{off}, and B was simulated as B^{on} and A^{off}. For the test simulations, uncovered reflects the activation of cues A^{on} and B^{off} (i.e. the condition where stimulus A was presented and the light corresponding to stimulus B was present but unilluminated) and covered by activation of A^{on} alone (i.e. the condition where stimulus A was presented and an opaque cover was placed over the light corresponding to stimulus B thus preventing access to that stimulus — that is, covering is implemented in the simulation by removing the B^{off} representation).

To summarise, if it is assumed that the non-illumination of a stimulus light is represented by animals, then, for Pearce's configural model (1994), the representations of such absences will contribute to the learning of a patterning discrimination. Moreover, the removal of such absence cues (by preventing access to the unilluminated light) will result in a disruption of performance because these, once predictive, cues are not able to contribute. In contrast, the elemental Rescorla-Wagner model (1972) is largely unaffected by the inclusion of CS^{off} cues. It should be noted that the simulations presented in Figure 1 reflects the default alpha values for the ALTSim program. Although no effort was made to optimise the simulation by the systematic manipulation of parameters, reducing the salience of the CS^{off} representations relative the CS^{on} representations makes little difference to the overall pattern of results. In addition, this pattern of results is not unique to the

models described here. Simulations of other models covered by ALTSim (Harris, 2006; Wagner, 2003) produce the same general pattern of results as Pearce's model. Indeed, if it is assumed that there are unique cues corresponding to the combination of two or more elements (Whitlow & Wagner, 1972) then the simulations of the Rescorla-Wagner model also correspond to the general pattern for Pearce's model. That is, any model tested here that can account for negative patterning will also be sensitive to the removal of CS^{off} cues.

In short, if it is assumed that the performance of animals exposed to a negative patterning discrimination is controlled by configural processes, while those only exposed to positive patterning is controlled by elemental learning mechanisms, then standard associative theories provide a perfect fit to the observed effects of preventing access to a cue during testing following training on patterning discriminations simply by considering the signal value of an unilluminated bulb. Thus there is no need to posit an additional process whereby the inability to access the source of a cue results in the status of that cue being treated as ambiguous. One might object that the assumption of different learning mechanisms for positive and negative patterning is entirely post-hoc. However, the same assumption was made by Fast and Blaisdell (in press), and this assumption is entirely consistent with the fact that positive patterning is typically acquired more readily than negative patterning (for recent demonstrations see, Harris, Gharaei, & Moore, 2009; Harris, Livesey, Gharaei, & Westbrook, 2008). Moreover, a consideration of CS^{off} cues can also provide an account of the sensory preconditioning experiments which instigated this line of research (Blaisdell et al., 2009). Consider that in those experiments A was paired with B before B was paired with food. As the B-food training was given across multiple trials, separated by intervals of between 2-6 minutes, there would have been the opportunity to learn an inhibitory relationship whereby B^{off} predicts the absence of food. Covering the stimulus light corresponding to B during test would mean that the B^{off} representation would remain inactivated and remove its inhibitory effects.

Although studies of patterning and sensory preconditioning have shown that rats are sensitive to the difference between a light being unilluminated and the covering of that stimulus bulb, the behaviour in these tests can be explained in terms of standard associative theory. To appropriate a phrase used judiciously by Fast and Blaisdell (in press) it is "as if" associative models are sensitive to the ambiguity created by the covering of a stimulus light. Obviously, these models are not actually sensitive to ambiguity as their behaviour is entirely specified by the formation of associative links between the stimulus representations it is furnished with. Entirely deterministic systems that have no way of representing or appreciating ambiguity produce outcomes that are functionally equivalent to that which might be produced by agents that were genuinely aware of ambiguity. Thus the functional outcome of behaviour is no guide to the mechanisms involved in its production. To take a florid example, one might describe a street magician with a set of marked cards as behaving "as if" they can read the minds of their audience, but this is hardly a proof of the existence of magic! Of course, there are many good reasons to pursue a functional analysis in a variety of situations (for a discussion of this issue in the context of rationality more broadly see Kacelnik, 2006) but

Morgan's Canon is explicitly concerned with the mechanisms by which a particular behaviour is produced rather than with its functional outcome.

Effort and Reward Value

We now move to an example, from our own work, of a pattern of behaviour that might seem to reflect the operation of complex, human-like, cognitive capacities. The value of a liquid reward can be assessed in rats by the examination of palatability via the microstructural analysis of licking during ingestion. Rats rarely show continuous consumption of a liquid. Instead they perform repeated clusters of licks separated by pauses. The mean number of licks per cluster has a positive, monotonic relationship with the concentration of a palatable solution such as sucrose (Davis & Smith, 1992; Spector, Klumpp, & Kaplan, 1998). Cluster size has also been found to have a negative monotonic relationship with the concentration of an unpalatable solution such as quinine (Hsiao & Fan, 1993; Spector & St. John, 1998). These findings have led to the idea that cluster size may be a useful measure of reinforcer palatability and hence, reward value. We have previously used this technique in a number of appetitive and aversive conditioning procedures (e.g., Dwyer, 2008, 2009; Dwyer, Boakes, & Hayward, 2008; Dwyer, Pincham, Thein, & Harris, 2009) but neither we, nor anyone else, had applied it in the context of instrumental behaviours. In the preliminary stages of a pilot experiment aimed at developing procedures in which rats would reliably press a lever to gain access to a sucrose solution, at the same time as their hedonic response to the sucrose reward could be measured, rats received exposure to increasing response requirements. Interestingly, it was observed that the size of lick clusters during ingestion of the sucrose reinforcer increased as the response requirement was raised. This increase in cluster size might be reflective of an effect whereby the more effort rats expended to obtain the reinforcer, the greater its perceived palatability. We (Lydall, Gilmour, & Dwyer, 2010) investigated this possibility by examining the value of the sucrose reward (using lick cluster size as an indicator of hedonic value) as a function of whether the rats were required to press a lever 50 times (high effort) to receive a sucrose reinforcer or whether they pressed a lever only 10 times (low effort). The effort involved in the lever press response requirement has two obvious components: work and waiting time, because it takes more work and more time to press a lever 50 times than 10. Therefore, that study also investigated the possibility that differences in the intervals between reinforcers in low and high ratio schedules may account for these changes in reward palatability using a procedure in which rats either worked (master conditions), or just waited to be rewarded (yoked conditions where the reinforcer occurred at times matched to those earned in the master conditions). The results indicated that rats perceived the sucrose reinforcer as more palatable when they pressed the lever 50 times than when they pressed 10 times. Although there was also a difference in the yoked conditions, this was smaller than in the master conditions, indicating that effort genuinely influenced the hedonic response to the reward.

This effect bears striking resemblance to the human phenomenon of 'Effort Justification,' in which greater value is placed on rewards that are harder to

obtain. According to Festinger (1957), this tendency can be attributed to cognitive dissonance: a state of psychological discomfort that occurs when there is a discrepancy between a person's attitude and their actions. In such circumstances, humans may try to reduce this dissonance by modifying their attitudes to complement their behaviour. In a classic study by Aronson and Mills (1959), young women underwent 'severe' or 'mild' initiation to join a discussion group. Participants in the severe initiation condition were asked to read aloud a sexually explicit passage, whereas those in the mild initiation group were given much less embarrassing material to read. When asked to rate the discussion group, the severe initiation group's ratings were significantly higher than those of the mild initiation group. That is, manipulating the effort required to join the group changed the participants' perceived value of the outcome of their behaviour. Aronson and Mills suggested that the group whose initiation process was more difficult increased their valuation of the discussion group to resolve the dissonant relationship produced by undergoing severe embarrassment in order to join a discussion group of little value.

Our results demonstrated that the value of a reward was indeed higher when it followed high rather than low effort, and such an effort justification effect in humans would likely be attributed to mechanisms such as cognitive dissonance. Indeed, the suggestion that cognitive dissonance might explain aspects of animal behaviour was championed by Lawrence and Festinger (1962). While the behavioural data they considered can be explained in far simpler terms (Mackintosh, 1974), this critique (largely focused on the fact that resistance to extinction is not positively related to the value of the reinforcer used in training) does not apply to our data. Therefore the analysis of effort effects in terms of cognitive dissonance offered by Lawrence and Festinger (1962) remains viable. But are there other explanations that might be offered that do not require rats to be subject to cognitive dissonance?

In an experiment by Clement, Feltus, Kaiser, and Zentall (2000), pigeons were required to peck a white light either once or twenty times to receive one of two different simultaneous colour discrimination problems. Pecks on the correct colour of each discrimination (S+) were rewarded with food. Following training, when pigeons were given a choice between the two positive (S+) colours, they showed a preference for the colour that followed the greatest effort during training. Clement et al. (2000) argued that their results could be attributed to 'within-trial contrast.' This account suggests that the value of a positive stimulus is judged relative to the motivational state of the animal prior to its presentation. A twenty-peck response requirement is presumably more aversive than a single peck requirement. Therefore the appearance of the S+ that normally follows twenty pecks may be a relatively greater improvement in conditions (a greater contrast) than the appearance of the S+ that normally follows one peck. In this way, a preference may develop for the outcome that follows higher effort. The experiment by Clement et al. is not a direct replication of the human effort justification effect, since the value of the reinforcer itself (the food reward) was not measured. However, the theoretical interpretation offered for their findings is interesting, especially as we have recently demonstrated that other forms of

contrast can produce changes in our lick-cluster measure of reward value (Dwyer, Lydall, & Hayward, 2011).

If it is assumed that high effort is more aversive than low effort then within-trial contrast provides a simple account of the current data. Moreover, if it is assumed that waiting a long time for reward is more aversive than waiting for a short time, the contrast account provides an explanation for the data in the yoked conditions as well. However, a consideration of the effects of the interval between reinforcers might offer an even simpler explanation. Each access to sucrose will lead to an adaptation of the relevant receptors, and our studies of simultaneous contrast demonstrate that adaptation will reduce the response to a given sucrose concentration (Dwyer et al., 2011). As recovery from adaptation will be larger with longer intervals, then such adaptation effects provide a direct account of the pattern of behaviour observed in the yoked conditions. Indeed, it is possible to extend this adaptation account to the high versus low effort comparison if lever pressing speeds recovery from adaptation. Nothing in the experiments considered above allows for a principled decision between the accounts in terms of contrast or adaptation but both of these accounts rest on far simpler mechanisms than those assumed by cognitive dissonance. Thus, although our results (Lydall et al., 2010) and those of others (e.g., Clement et al., 2000) might appear on the surface to support Lawrence and Festinger's (1962) contention that cognitive dissonance contributes to the behaviour on non-human animals the existence of simpler accounts for that behaviour renders this interpretation unnecessary.

What is "Simple"?

At times in the preceding discussions we have considered models of associative learning that are, to a greater or lesser degree, computationally complex enough that deriving predictions from them requires a process of computer-based simulation. This was most striking when comparing the causal model theory account of Blaisdell et al.'s (2006) experiments with simulations of the Schmajuk et al. (1996) model: the predictions of causal model theory can be derived easily and directly from the theoretical account itself, while the Schmajuk et al. (1996) model requires detailed simulation to derive its predictions. Thus, although an account in terms of causal reasoning requires accepting that rats are capable of higher-order mental processes, it is (computationally at least) simpler than an account in terms of lower-order (associative/mechanistic) processes. As such, two different conceptions of simplicity appear to be in tension here: Morgan's Canon explicitly favours accepting the explanation in terms of the lowest form of mental process, while a broader conception of parsimony (as encapsulated by Occam's Razor) would favour the explanation that requires the lower number of postulated processes regardless of whether they correspond to higher or lower-order mental processes.

Morgan (1894) himself scathingly dismissed the idea that apparently simpler explanations might be preferable without a consideration of the level of mental process involved:

Is it not simpler to explain the higher activities of animals as the direct outcome of reason or intellectual thought? … It is the apparent simplicity of the explanation that leads many people to naively adopt it. But surely the simplicity of an explanation is no necessary criterion of its truth. The explanation of the genesis of the organic world by direct creative fiat is far simpler than the explanation of its genesis through the indirect method of evolution. (p. 54)

Critically, attributing human-like levels of mental processes to animals only allows for superficially simpler accounts of their behaviour. The vast literature on, for example, reasoning and inference in humans is eloquent testimony to the fact that explaining the cognitive processes underpinning such terms is incredibly complex in its own right. In short, the apparent contrast between Morgan's Canon, with its emphasis on the level of mental processes, and more general conceptions of parsimony rests, at least here, on a mischaracterisation of higher mental processes as "simple" conceptually. In the present case, the "higher-order" processes being considered such as causal reasoning themselves require further explanation before they can truly be said to be understood, whereas the "lower" associative level processes can be described in purely mechanistic terms, and thus do not require further explanation (the details of their biological instantiation aside). This is not to say that describing behaviour in terms of higher-order mental processes is invalid or incorrect, merely that the conceptual complexity of these processes should not be ignored.

Summary and Conclusions

We have considered three areas, causal-reasoning, sensitivity to the absence of stimuli, and the relationship between effort and reward, where higher-order mental processes might be invoked as explanations for the behaviour of rats. We have deliberately not explored whether any of the higher-order accounts are directly inconsistent with the observed data (although this is not to say that these accounts cannot, or should not, be tested in such a fashion) because it is our intention to highlight the principle enshrined in Morgan's Canon; namely that accounts of animal behaviour in terms of higher-order mental functions should only be accepted when explanations in terms of simpler mechanisms are unavailable. That is, accepting Morgan's Canon places the onus of proof on those that would challenge the adequacy of conceptually simple accounts of non-human behaviour. In each of the cases considered here we have discussed alternative accounts of those behaviours that did not rely on such higher-order processes and argued that these simpler mechanisms are consistent with the observed data. Therefore, if Morgan's Canon still represents a grounding assumption for comparative psychology, the lesson in all cases is the same: despite surface appearances, none of the evidence we have discussed here supports the conclusion that the mental capacities of rats include higher-order rational processes.

References

Aronson, E., & Mills, J. (1959). The effect of severity of initiation on liking for a group. *Journal of Abnormal and Social Psychology, 59*, 177-181.

Blaisdell, A. P., Leising, K. J., Stahlman, W., & Waldmann, M. R. (2009). Rats distinguish between absence of events and lack of information in sensory preconditioning. *International Journal of Comparative Psychology, 22*, 1-18.

Blaisdell, A. P., Sawa, K., Leising, K. J., & Waldmann, M. R. (2006). Causal reasoning in rats. *Science, 311*, 1020-1022.

Clayton, N., & Dickinson, A. (2006). Rational rats. *Nature Neuroscience, 9*, 472-474.

Clement, T. S., Feltus, J. R., Kaiser, D. N., & Zentall, T. R. (2000). "Work ethic" in pigeons: Reward value is directly related to the effort or time required to obtain the reward. *Psychonomic Bulletin & Review, 7*, 100-106.

Davis, J. D., & Smith, G. P. (1992). Analysis of the microstructure of the rhythmic tongue movements of rats ingesting maltose and sucrose solutions. *Behavioral Neuroscience, 106*, 217-228.

Dickinson, A., & Burke, J. (1996). Within-compound associations mediate the retrospective revaluation of causality judgments. *Quarterly Journal of Experimental Psychology, 49B*, 60-80.

Dwyer, D. M. (2003). Learning about cues in their absence: Evidence from flavour preferences and aversions. *Quarterly Journal of Experimental Psychology, 56B*, 56-67.

Dwyer, D. M. (2008). Microstructural analysis of conditioned and unconditioned responses to maltodextrin. *Learning & Behavior, 36*, 149-158.

Dwyer, D. M. (2009). Microstructural analysis of ingestive behaviour reveals no contribution of palatability to the incomplete extinction of a conditioned taste aversion. *Quarterly Journal of Experimental Psychology, 62*, 9-17.

Dwyer, D. M., Boakes, R. A., & Hayward, A. J. (2008). Reduced palatability in lithium- and activity-based, but not in amphetamine-based, taste aversion learning. *Behavioral Neuroscience, 122*, 1051-1060.

Dwyer, D. M., Lydall, E. S., & Hayward, A. J. (2011). Simultaneous contrast: Evidence from licking microstructure and cross-solution comparisons. *Journal of Experimental Psychology: Animal Behavior Processes, 37*, 200-210.

Dwyer, D. M., Pincham, H. L., Thein, T., & Harris, J. A. (2009). A learned flavor preference persists despite the extinction of conditioned hedonic reactions to the cue flavors. *Learning & Behavior, 37*, 305-310.

Dwyer, D. M., Starns, J., & Honey, R. C. (2009). "Causal reasoning" in rats: A reappraisal. *Journal of Experimental Psychology: Animal Behavior Processes, 35*, 578-586.

Fast, C. D., & Blaisdell, A. P. (in press). Rats are sensitive to ambiguity. *Psychonomic Bulletin & Review*. doi: 10.3758/s13423-011-0171-0

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.

Foote, A. L., & Crystal, J. D. (2007). Metacognition in the rat. *Current Biology, 17*, 551-555.

Fujita, K., Kuroshima, H., & Asai, S. (2003). How do tufted capuchin monkeys (*Cebus apella*) understand causality involved in tool use? *Journal of Experimental Psychology: Animal Behavior Processes, 29*, 233-242.

Hall, G., & Honey, R. C. (1989). Contextual effects in conditioning, latent inhibition, and habituation — associative and retrieval functions of contextual cues. *Journal of Experimental Psychology: Animal Behavior Processes, 15*, 232-241.

Harris, J. A. (2006). Elemental representations of stimuli in associative learning. *Psychological Review, 113*, 584-605.

Harris, J. A., Gharaei, S., & Moore, C. A. (2009). Representations of single and compound stimuli in negative and positive patterning. *Learning & Behavior, 37*, 230-245.

Harris, J. A., Livesey, E. J., Gharaei, S., & Westbrook, R. F. (2008). Negative patterning is easier than a biconditional discrimination. *Journal of Experimental Psychology: Animal Behavior Processes, 34*, 494-500.

Holland, P. C. (1990). Event representation in Pavlovian conditioning: Image and action. *Cognition, 37*, 105-131.

Hsiao, S., & Fan, R. J. (1993). Additivity of taste-specific effects of sucrose and quinine — microstructural analysis of ingestive behavior in rats. *Behavioral Neuroscience, 107*, 317-326.

Kacelnik, A. (2006). Meanings of rationality. In S. Hurley & M. Nudds (Eds.), *Rational animals?* (pp. 87-106). New York: Oxford University Press.

Kutlu, M. G., & Schmajuk, N. A. (in press). Classical conditioning mechanisms can differentiate between seeing and doing in rats. *Journal of Experimental Psychology: Animal Behavior Processes*.

Lawrence, D. H., & Festinger, L. (1962). *Deterrents and reinforcement: The psychology of insufficient reward*. Stanford, CA: Stanford University Press.

Leising, K. J., Wong, J., Waldmann, M. R., & Blaisdell, A. P. (2008). The special status of actions in causal reasoning in rats. *Journal of Experimental Psychology: General, 137*, 514-527.

Lydall, E. S., Gilmour, G., & Dwyer, D. M. (2010). Rats place greater value on rewards produced by high effort: An animal analogue of the "effort justification" effect. *Journal of Experimental Social Psychology, 46*, 1134–1137.

Mackintosh, N. J. (1974). *The psychology of animal learning*. Oxford, England: Academic Press.

Martin-Ordas, G., Call, J., & Colmenares, F. (2008). How great apes perform in functionally equivalent trap-tube and trap-table tasks. *Folia Primatologica, 79*, 359-360.

Morgan, C. L. (1894). *An introduction to comparative psychology*. London, England: Walter Scott Publishing Co.

Murphy, R. A., Mondragon, E., & Murphy, V. A. (2008). Rule learning by rats. *Science, 319*, 1849-1851.

Pearce, J. M. (1994). Similarity and discrimination - a selective review and a connectionist model. *Psychological Review, 101*, 587-607.

Rescorla, R. A. (1980). Simultaneous and successive associations in sensory preconditioning. *Journal of Experimental Psychology: Animal Behavior Processes, 6*, 207-216.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current theory and research* (pp. 64-99). New York: Appelton-Century-Crofts.

Rizley, R. C., & Rescorla, R. A. (1972). Associations in second-order conditioning and sensory preconditioning. *Journal of Comparative and Physiological Psychology, 81*, 1-11.

Santos, L. R., Pearson, H. M., Spaepen, G. M., Tsao, F., & Hauser, M. D. (2006). Probing the limits of tool competence: Experiments with two non-tool-using species (*Cercopithecus aethiops* and *Saguinus oedipus*). *Animal Cognition, 9*, 94-109.

Schmajuk, N. A., Lam, Y. W., & Gray, J. A. (1996). Latent inhibition: A neural network approach. *Journal of Experimental Psychology: Animal Behavior Processes, 22*, 321-349.

Spector, A. C., Klumpp, P. A., & Kaplan, J. M. (1998). Analytical issues in the evaluation of food deprivation and sucrose concentration effects on the microstructure of licking behavior in the rat. *Behavioral Neuroscience, 112*, 678-694.

Spector, A. C., & St. John, S. J. (1998). Role of taste in the microstructure of quinine ingestion by rats. *American Journal of Physiology-Regulatory Integrative and Comparative Physiology, 274*, 1687-1703.

Tebbich, S., & Bshary, R. (2004). Cognitive abilities related to tool use in the woodpecker finch, *Cactospiza pallida*. *Animal Behaviour, 67*, 689-697.

Thorwart, A., Schultheis, H., Konig, S., & Lachnit, H. (2009). ALTSim: A MATLAB simulator for current associative learning theories. *Behavior Research Methods, 41*, 29-34.

Visalberghi, E., & Limongelli, L. (1994). Lack of comprehension of cause-effect relations in tool-using capuchin monkeys (*Cebus apella*). *Journal of Comparative Psychology, 108*, 15-22.

Wagner, A. R. (2003). Context-sensitive elemental theory. *Quarterly Journal of Experimental Psychology, 56B*, 7-29.

Waldmann, M. R., & Hagmayer, Y. (2005). Seeing versus doing: Two modes of accessing causal knowledge. *Journal of Experimental Psychology: Learning Memory and Cognition, 31*, 216-227.

Waldmann, M. R., Hagmayer, Y., & Blaisdell, A. P. (2006). Beyond the information given: Causal models in learning and reasoning. *Current Directions in Psychological Science, 15*, 307-311.

Whitlow, J. W., & Wagner, A. R. (1972). Negative patterning in classical conditioning: Summation of response tendencies to isolable and configural components. *Psychonomic Science, 27*, 299-301.