

Journal Club

Editor's Note: These short, critical reviews of recent papers in the *Journal*, written exclusively by graduate students or postdoctoral fellows, are intended to summarize the important findings of the paper and provide additional insight and commentary. For more information on the format and purpose of the Journal Club, please see http://www.jneurosci.org/misc/ifa_features.shtml.

Multiple Interpretations of Cocaine-Seeking Behavior after Prolonged Self-Administration Training

Nathan M. Holmes^{1*} and Kelly J. Clemens^{2*}

¹School of Psychology, University of New South Wales, Sydney, New South Wales 2052, Australia, and ²Department of Psychology, Macquarie University, Sydney, New South Wales 2109, Australia

Review of Zapata et al.

In the past 30 years, seminal research has identified a dual system of instrumental response control: a goal-directed [or response–outcome (R–O)] system that permits responses to be deployed in the service of current needs and desires, and a habitual (or stimulus–response) system that permits environmental (or discriminative) stimuli to automatically elicit responses independently of needs and desires. These two systems are said to compete with one another for response control. Early in operant training, the goal-directed system is dominant and instrumental responses are sensitive to changes in both the R–O contingency (Ballesine and Dickinson, 1998) and outcome value (Adams and Dickinson, 1981). However, with repeated training, the habitual system becomes increasingly dominant such that changes in R–O contingency and outcome value have a reduced impact on instrumental performance (Dickinson, 1985; Dickinson et al., 1995). Thus, it appears that there is a transition from internal voluntary control to external stimulus-driven control with increasing amounts of training. This transition is adaptive in the sense that performance can be maintained while the load on cognitive resources is reduced.

A transition from voluntary, controlled drug taking toward a loss of control and compulsive drug intake is a hallmark of addiction. Modeling this behavior is an important step toward understanding the neural processes by which drug seeking persists, often in the face of adverse consequences. The standard approach has been to focus on extending the findings on instrumental response control using natural rewards to intravenously administered drug rewards. However, this enterprise has met with limited success, owing largely to the fact that standard methods of reward devaluation (such as lithium-induced nausea or specific satiety) are ineffective with a reward that has no gustatory or consummatory component.

In a recent article published in the *Journal of Neuroscience*, Zapata et al. (2010) attempt to circumvent this problem. In this study, rats were trained to perform a drug seeking/taking chained schedule to obtain a cocaine reward. Here, responding on the first “drug-seeking” lever resulted in the retraction of this lever and insertion of a second “drug-taking” lever. Responding on this second lever led to a single infusion of cocaine. Both levers remained retracted during a 600 s time-out period before the drug-seeking lever was once more available on a random interval schedule (120 s). At the conclusion of either limited (minimum six sessions, two sessions/d) or extended (minimum 36 sessions, two sessions/d) training on this chained schedule,

responding on the drug-taking lever was extinguished. During this phase, the drug-taking lever was presented and responses on this lever no longer resulted in a cocaine infusion. Rats were then tested for performance when only the drug-seeking lever was available. Zapata et al. (2010) found that sensitivity of the drug-seeking response to extinction of the drug-taking response differed depending on the amount of training; after limited training, rats withheld responding on the drug seeking lever, but after extended training, rats maintained responding on this lever. Zapata et al. (2010) reasoned that after limited training, drug-seeking responses continued to be goal-directed (where the goal is to obtain access to the drug-taking lever) and thus sensitive to extinction of the drug-taking response. After extended training, drug-seeking responses had become habitual and thus insensitive to extinction of the drug-taking response. In a second experiment, the training procedure was repeated but rats received bilateral infusions of lidocaine into the dorsolateral striatum (DLS) before test. Inactivation of the DLS in extensively trained rats reduced responding on the drug-seeking lever at test, returning it to a level similar to that seen in rats with only limited training. The authors take these findings as evidence that drug-seeking responses become less sensitive to extinction of the drug-taking response, and thus habitual, across extended training; and that inactivation of the DLS can reinstate

Received Dec. 6, 2010; revised Dec. 19, 2010; accepted Dec. 23, 2010.

*N.M.H. and K.J.C. contributed equally to this work.

Correspondence should be addressed to Kelly J. Clemens, Neuropharmacology Laboratory, Department of Psychology, C3A, Macquarie University, NSW 2109, Australia. E-mail: kelly.clemens@mq.edu.au.

DOI:10.1523/JNEUROSCI.6354-10.2011

Copyright © 2011 the authors 0270-6474/11/313935-02\$15.00/0

goal-directed control over drug-seeking performance.

Although Zapata et al. (2010) have clearly demonstrated that there are indeed important differences between responses performed early versus late in training, caution should be taken in attributing the persistence of drug-seeking after extended training to a loss of voluntary response control. Although the authors discuss reward devaluation, it must be noted that cocaine itself is never devalued, nor associated with any adverse consequence. This leaves open the possibility that, after extended training, rats continue to perform the drug-seeking response for some reason other than because it has become a habit.

One alternative explanation for the persistence of drug seeking after extended training in the Zapata et al. (2010) study is that rats developed a direct association between the initial response and the outcome of the chain. In support of this idea, Corbit and Balleine (2003) have shown that after rats learn to perform a sequence of left and right lever presses to obtain a food pellet reward, devaluation of the reward by prefeeding to satiety suppressed performance of the first but not the second response in the chain. This selective effect of devaluation on the first (seeking) response can only have resulted from a direct association between this response and the outcome of the sequence. This is especially relevant to the Zapata et al. (2010) study because, as noted above, cocaine itself was never devalued. Thus, with extended training, persistence of the drug-seeking response after extinction of the drug-taking response may have been simply due to a direct association between the drug-seeking response and cocaine, rather than a loss of control over cocaine-seeking responses.

A second alternative explanation for the persistence of drug seeking after extended training is that rats come to represent the task differently across these sessions. For a long time, it has been assumed that when rats are trained to perform a chain of two responses to procure a reward, performance of the initial response (i.e., the distal or drug-seeking response) is maintained by the conditioned reinforcing properties of the second re-

sponse (i.e., the proximal or drug-taking response—the response associated with the primary reward). Early in training, this may well be the case. However, in a study using natural reward and the same heterogeneous chain used by Zapata et al. (2010), Ostlund et al. (2009) demonstrated that across training, each response in the sequence ceases to be represented as a discrete behavioral unit. Instead, the two responses in the sequence became “chunked” into a single representation that itself is capable of entering into associations. Given the amount of training that rats received in the Zapata et al. (2010) study, chunking of the drug-seeking and drug-taking responses may have occurred. After chunking, there is no reason why extinction of the drug-taking response should affect performance of the drug-seeking response. As the outcome of the seeking–taking action sequence remains highly valued, it is not surprising that rats continue initiating the sequence in pursuit of this outcome. Again, this explanation for the persistence of the drug-seeking response after extended training in the Zapata et al. (2010) study provides an alternative interpretation of the authors’ claim that drug-seeking responses had become habitual across this training.

The above ideas can explain the behavioral results presented in the Zapata et al. (2010) study (experiment 1). But what of their finding that DLS inactivation restores sensitivity of drug seeking to extinction of the drug-taking response (experiment 2)? Of critical importance to this question, Ostlund et al. (2009) additionally investigated the neural substrates of action sequence chunking. Whereas rats with sham lesions withheld performance of a two-step chain when the outcome of that sequence was devalued, rats with lesions of the medial agranular cortex showed differential performance on the two levers; they continued to perform the first response but ceased to perform the second. Ostlund et al. (2009) argued that lesioned rats failed to chunk the two responses into a single representation and instead represented each response as a discrete behavioral unit, performing the task in a chain-like manner. Moreover, Ostlund et al. (2009) noted that there are rich connections between the medial agranular cortex and the

dorsal striatum. This may provide an alternative explanation as to why rats with DLS inactivation in the Zapata et al. (2010) study withheld drug-seeking responses after extended training and extinction of the drug-taking response. Like Ostlund et al.’s (2009) lesioned rats, DLS inactivation may have caused Zapata et al.’s (2010) rats to revert to chain-like performance. Hence, their drug-seeking performance was again sensitive to extinction of the drug-taking response.

In summary, persistence of drug seeking after extended training and extinction of the drug-taking response may be supported by a direct association between the drug-seeking response and cocaine or, alternatively, through the entire sequence of drug-seeking and drug-taking responses becoming associated with cocaine. Thus, it is not clear if the persistence of drug seeking under these conditions results from insensitivity of these responses to reward value. In the same vein, the finding that DLS inactivation reduces drug-seeking responses after extended training and extinction of the drug-taking response does not necessarily imply a restoration of goal-directed control over performance. This result may simply reflect that DLS rats had reverted to a chain-like mode of performance.

References

- Adams CD, Dickinson (1981) Instrumental responding following reinforcer devaluation. *J Exp Psychol B* 33:109–121.
- Balleine BW, Dickinson A (1998) Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology* 37:407–419.
- Corbit LH, Balleine BW (2003) Instrumental and Pavlovian incentive processes have dissociable effects on components of a heterogeneous instrumental chain. *J Exp Psychol Anim Behav Process* 29:99–106.
- Dickinson A (1985) Actions and habits: the development of behavioural autonomy. *Philos Trans R Soc Lond B* 308:67–78.
- Dickinson A, Balleine BW, Watt A, Gonzalez F, Boakes RA (1995) Motivational control after extended instrumental training. *Anim Learn Behav* 23:197–206.
- Ostlund SB, Winterbauer NE, Balleine BW (2009) Evidence of action sequence chunking in goal-directed instrumental conditioning and its dependence on the dorsomedial prefrontal cortex. *J Neurosci* 29:8280–8287.
- Zapata A, Minney VL, Shippenberg TS (2010) Shift from goal-directed to habitual cocaine seeking after prolonged experience in rats. *J Neurosci* 30:15457–15463.