Exploring the relationship between actions, habits, and automaticity in an action sequence task

Eric Garr[1,2] and Andrew R. Delamater[1,2]

[1]*Graduate Center, City University of New York, New York 10016, USA;* [2]*Brooklyn College, City University of New York, New York 11210, USA*

It is tempting to equate the automatization of an action sequence with the formation of a habit. However, the term "habit" specifically implies a failure to evaluate future consequences to guide behavior. To test if automatized sequences become habitual, we trained rats on an action sequence task for either 20 or 60 d and then conducted reward devaluation tests. While both groups showed equivalent goal-directed performance of the trained action sequence on a global measure of behavior, sequence initiation and completion times were differentially sensitive to outcome devaluation in moderately and extensively trained rats.

[Supplemental material is available for this article.]

It is thought that extensive practice creates habits. What is meant by a habit, however, is not always clear. On the one hand, an action can be habitual in the sense that it is insensitive to the anticipated value of its consequence (e.g., Dickinson et al. 1983). On the other hand, the term "habitual" is used to refer to sequences of actions that are performed with a high level of automaticity—that is, high speed and low variability (Jog et al. 1999; Desrochers et al. 2015). This raises the question of whether automatized action sequences are controlled by an anticipation of future outcomes.

According to one view (Dezfouli and Balleine 2013), well-learned action sequences are habitual in that the component parts are executed without evaluating future outcomes. However, prior to the sequence being executed, the subject is hypothesized to engage in goal-directed decision-making such that the initiation of a sequence is subject to goal-directed control. This theory predicts that, when the outcome of a sequence is devalued, subjects should perform few sequences because they will be less prone to initiate them. However, on the rare occasion that a sequence is initiated, the component actions that comprise the sequence should be performed rapidly regardless of the value of the outcome.

To test these predictions, we performed an experiment by using an outcome devaluation procedure after rats were given either moderate or extensive training on an action sequence task. While it has been demonstrated that there is a transition from goal-directed to habitual control with overtraining in single-response tasks (e.g., Adams 1982), it is unclear how subjects sequence their actions under these conditions. Other multiple-response tasks have been designed to study how outcome devaluation affects sequence performance (Balleine et al. 2005; Ostlund et al. 2009), but the free-operant nature of these tasks may hinder the development of automaticity because these tasks allow many different sequences to be reinforced. This is not ideal if the goal is to elicit repetitive behavior. In contrast, our discrete-trial task explicitly reinforces a clearly defined action sequence, and permits a more thorough exploration of the role of automaticity in action control.

Sixty-four naïve Long–Evans rats ($n = 32$ for each of two replications) were maintained at 85% of their ad libitum weight for the duration of the experiment, with water freely available throughout. Rats were first given magazine training in operant chambers

(MED Associates) with one pellet type (TestDiet MLabRodent 45 mg grain or Bio-serv 45 mg purified pellets, counterbalanced across rats). During this 20-min session, pellets were delivered according to a 60 sec random time schedule, and accompanied by a brief clicker (15 Hz for 0.5 sec).

Rats were then trained to press levers. During the first session of pretraining, the left lever was inserted. A press on the left lever resulted in pellet delivery into the magazine, the retraction of the left lever, and insertion of the right lever. A press on the right lever resulted in pellet delivery into the magazine, the retraction of the right lever, and insertion of the left lever. This cycle continued until 50 pellets were earned or 60 min elapsed, whichever occurred first. Three rats did not learn to press the levers during this phase of training and were excluded from the remainder of the experiment. A second pretraining session was given 24 h later, in which the conditions were identical to the previous session except that pellets were only delivered following a right lever press. The main training phase began 24 h later and continued for either 20 or 60 daily sessions (Fig. 1). During these sessions, the left and right levers were simultaneously inserted at the beginning of every trial, where they remained inserted until the rat completed a sequence of two lever presses. There were four possible sequences that could be performed: left–left (LL), left–right (LR), right–left (RL), or right–right (RR). If the rat performed an LR sequence, a pellet was delivered and the levers retracted for 1.5 sec before being inserted again to start the next trial. If the rat performed any other two-lever sequences pellets were not delivered and the levers retracted for 5 sec. A similar version of this task has been used previously with mice (Yin 2009, 2010; Rothwell et al. 2015). Sessions ended when 50 pellets were earned or 30 min elapsed, whichever occurred first. Six rats failed to learn the task and were excluded from the remainder of the experiment.

Two groups of rats were trained on the action sequence task. One group (moderate training; 14 males and 14 females) was trained for 20 daily sessions and another group (extensive training;
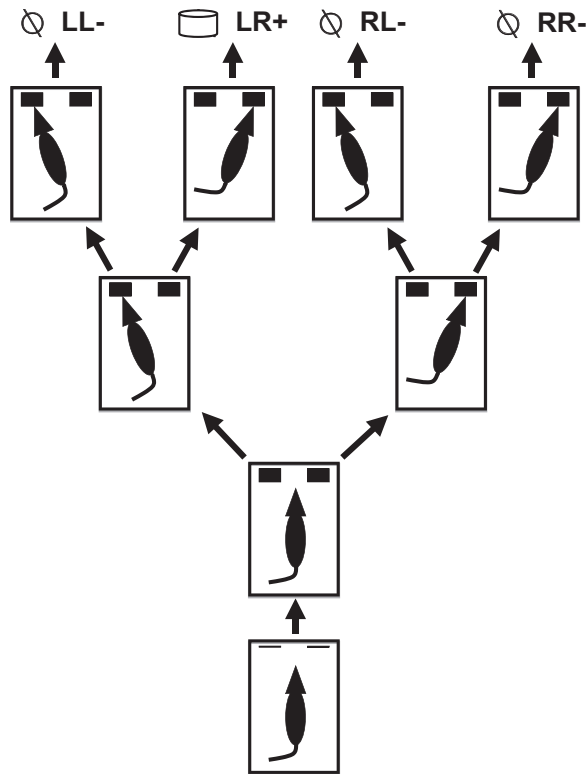
Figure 1. An illustration of the action sequence task.

12 males and 15 females) was trained for 60 daily sessions. The moderate group began training on the same day that the extensive group began day 41 of training so that both groups terminated training on the same day. Group assignment, pellet assignment, and sex were counterbalanced.

Following training, the devaluation cycles began. A single devaluation cycle was comprised of two tests separated by a retraining session, and all rats experienced two devaluation cycles. Prior to each test, rats were given an hour of unlimited access to either the pellet type associated with LR sequences (devalued test) or the other pellet used as a control for general satiety (valued test), with order of testing counterbalanced. All rats were preexposed to the novel pellet type 24 h prior to testing. Immediately after the satiation period, rats were placed in the operant chambers and given a 5-min extinction test in which the levers operated exactly as they did during training except no pellets were delivered and the clicker was turned off. Immediately following each extinction test, a 20-min preference test was conducted wherein rats were given a choice between the two pellet types to test the effectiveness of the selective satiety manipulation.

For statistical analysis, we performed t-tests to assess between-group differences on various training measures, between-group ANOVAs to assess differences in performance accuracy at the end of training , and one-way repeated measures ANOVAs (with pooled error terms across groups) to assess devaluation effects. Significant effects involving more than two means were assessed by constructing a set of v_1 mutually orthogonal post-hoc contrasts (Rodger 1974). This approach eliminates the interaction term from the linear model together with the problems associated with interaction tests (see Rodger 1974), and is more powerful than most ANOVA techniques at detecting true effects (Rodger and Roberts 2013). We also provide a measure of effect size based on Perlman and Rasmussen's (1975) estimate of the noncentrality parameter Δ.

Type I error is defined on a per contrast basis as the expected rate of rejecting true null hypotheses, and Rodger (1975) provided tables of critical F values for $α = 0.05$, the criterion adopted here.

Analysis of the training data revealed that extensively trained rats were more repetitive and, in some respects, less variable than moderately trained rats. On the last block of training, a one-way ANOVA (collapsed across groups) revealed significant differences among the mean proportion of the different sequence types ($F_{(3,159)} = 271.87$, MSE = 0.02, Δ = 802.35, $P < 0.05$). Post-hoc contrasts revealed that sequence frequency took the following ordering: LR > RR > LL > RL, indicating that both groups learned the task (Fig. 2A). However, extensively trained rats performed relatively more LR sequences and relatively fewer RR sequences compared to moderately trained rats ($Fs_{(1,160)} > 10.69$, MSE = 0.02, Δ = 9.56, $P < 0.05$).

To further characterize the sequence distributions, we calculated the normalized entropy for each rat (Fig. 2B). Normalized entropy, also known as U-value (Neuringer 2002), is calculated:

$$\frac{-\sum_{i=1}^{n} RF_i \times \log_2(RF_i)}{\log_2(n)},$$

where RF signifies the relative frequency of a sequence and n is the total number of possible sequences. If a rat behaves randomly, the expected distribution of sequences is uniform and the U-value is 1. If the rat performs only one sequence, then the U-value is 0. Although the U-values generally decreased across training in both groups, on the last block of training the moderate group was significantly greater than the extensive group ($t_{(53)} = 3.72$, $P < 0.05$). The moderate group also performed fewer correct sequences per minute compared to the extensive group (Fig. 2C; $t_{(53)} = 3.38$, $P < 0.05$).

We also measured the latency to initiate and complete correct sequences. Correct initiation latency was defined as the time separating insertion of the levers and a left lever press on LR trials (Fig. 2D). Initiation latency data during the training phase were only available from the second replication. By the last block of training, initiation times did not differ between groups ($t_{(24)} = 0.73$, $P > 0.05$). To calculate the variability in how quickly LR sequences were initiated, we calculated the coefficient of variation (CV) of the initiation times. The extensive group was less variable by the last training block ($t_{(24)} = 2.47$, $P < 0.05$). To calculate correct completion latency, the time separating a left lever press from a right lever press on LR trials was measured (Fig. 2E). On the last block of training, the two groups did not differ in their mean LR completion times ($t_{(53)} = 1.50$, $P > 0.05$), nor did they differ in LR completion time CV ($t_{(53)} = 0.65$, $P > 0.05$).

Next, we analyzed devaluation test data. We first examined the rate at which each sequence type was performed during valued and devalued tests (Fig. 3A). Separate one-way repeated measures ANOVAs revealed significant differences among the means for moderate ($F_{(7,371)} = 36.52$, MSE = 1.31, Δ = 247.26, $P < 0.05$) and extensive ($F_{(7,371)} = 71.23$, MSE = 1.31, Δ = 488.92, $P < 0.05$) groups. Post-hoc contrasts revealed devaluation effects only on LR sequences for each group (moderate: $F_{(7,371)} = 4.60$, $P < 0.05$; extensive: $F_{(7,371)} = 2.62$, $P < 0.05$). Additionally, the moderate group performed fewer sequences overall ($F_{(1,53)} = 5.07$, MSE = 1.94, Δ = 3.88, $P < 0.05$). It thus appears that truly extensive training on an action sequence task does not result in overall habitual performance.

We next examined initiation and completion latencies during these devaluation tests. If an action is governed by an anticipation of the outcome then devaluing the outcome should slow the time to respond. For sequence initiation times (i.e., the time separating lever insertion from a left lever press; Fig. 3B) the moderate group was slower on devalued than valued tests ($F_{(1,53)} = 5.54$, MSE =
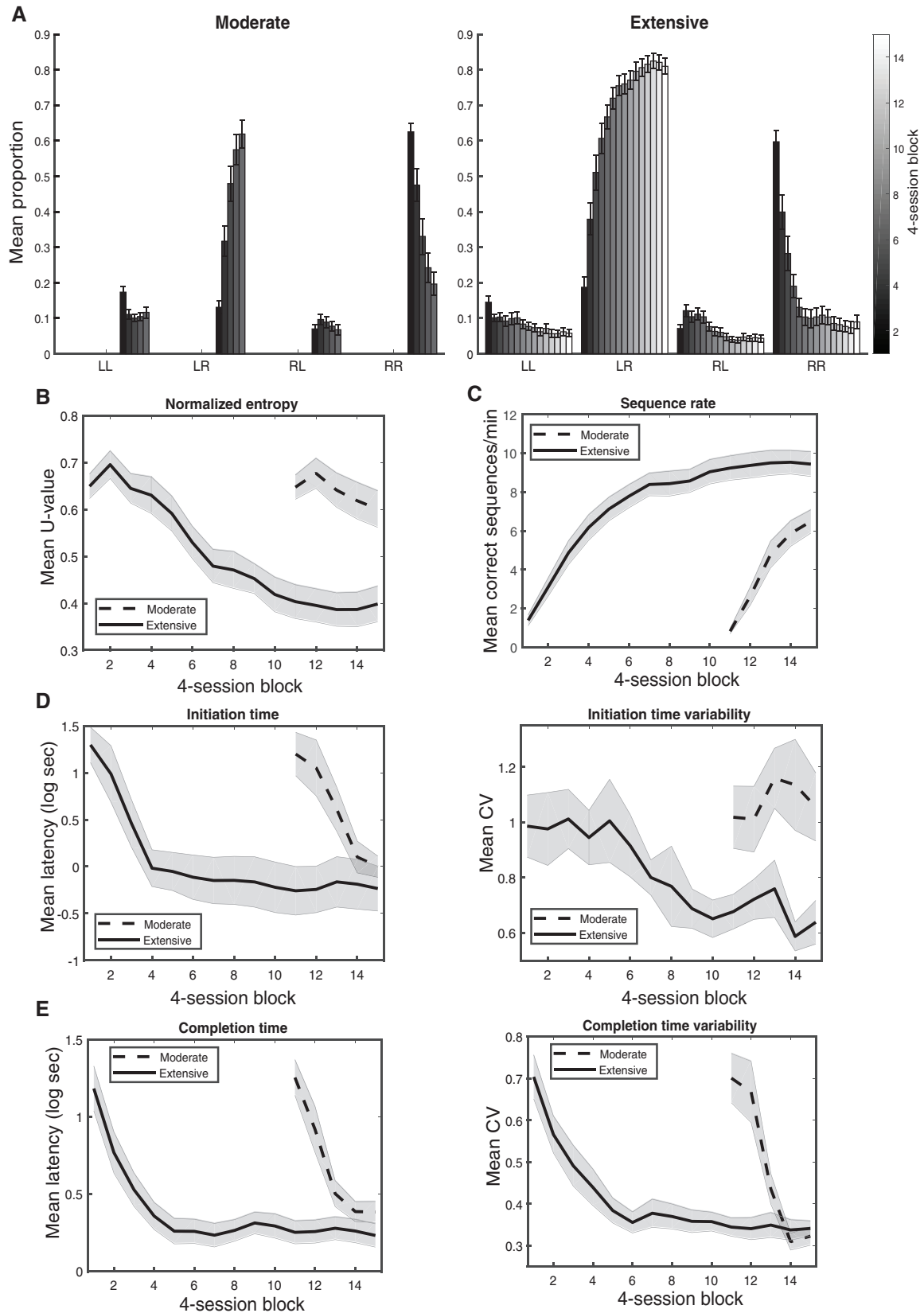
Figure 2. Training data. (*A*) Sequence distributions for each group as a function of four-session blocks. Error bars are ±SEM. (*B*) Normalized entropy (also termed U-value) as a function of four-session blocks. (*C*) The number of LR sequences per minute as a function of four-session blocks. (*D*) The mean time to initiate an LR sequence (*left*) and the mean variability with which LR sequences were initiated (*right*). Data come exclusively from the second replication. (*E*) Same as *D* but applied to completion of LR sequences. Light gray bounds are ±SEM.
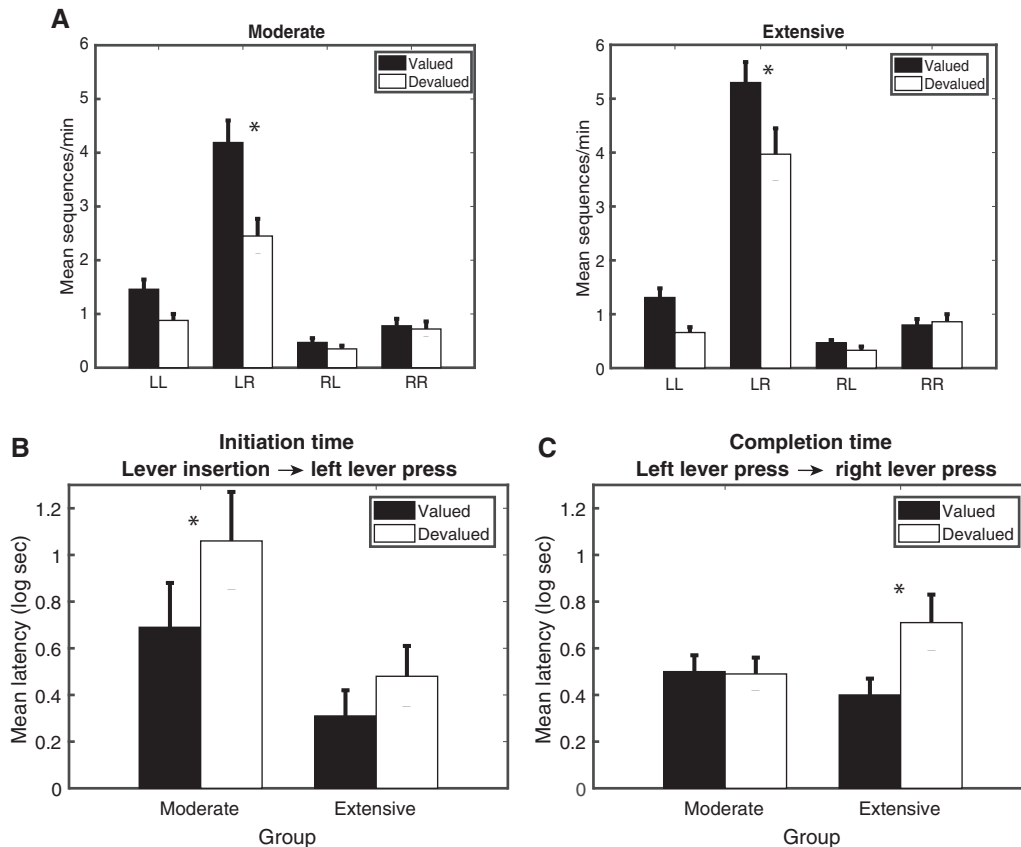
Figure 3. Devaluation test data. (*A*) Sequence distributions for each group as a function of valued and devalued test sessions. "Devalued" refers to when a rat was sated on the pellet type associated with LR sequences, and "valued" refers to being sated on the control pellet type. (*B*) Time to initiate a left lever press as a function of valued and devalued test sessions. (*C*) Time to complete LR sequences. One rat in the moderate group did not perform any LR sequences during the devalued test sessions, and thus did not contribute data to this graph. Error bars are ±SEM. (*) statistically significant difference.

0.34, $\Delta = 4.33$, $P < 0.05$), while the extensive group did not reliably show this difference ($F_{(1,53)} = 1.14$, $P > 0.05$). There was also a main effect of group, with the extensive group displaying overall faster initiation latencies ($F_{(1,53)} = 5.53$, MSE = 1.18, $\Delta = 4.32$, $P < 0.05$). Additional analyses revealed that nontarget initiation times were sensitive to devaluation (see Supplemental Material). For sequence completion times (i.e., the time from a left lever press to a right lever press; Fig. 3C) the extensive group was slower on devalued tests ($F_{(1,52)} = 8.68$, MSE = 0.09, $\Delta = 7.35$, $P < 0.05$), while the moderate group did not reliably show this difference ($F_{(1,52)} = 0.39$, $P > 0.05$). There was no main effect of group ($F_{(1,52)} = 0.15$, $P > 0.05$). Additional analyses of nontarget completion times revealed a devaluation effect only for RL sequences in extensively trained rats (see Supplemental Material). Extensively trained rats were also faster than moderately trained rats to entered the food magazine following an LR sequence, but magazine entry times were devaluation insensitive (see Supplemental Material). Thus, it appears that the extent of training determines where in the sequence goal-directed control manifests itself, with moderately trained rats showing greater hesitation to initiate and extensively trained rats showing greater hesitation to complete following reward devaluation.

Consumption data from the satiation periods indicate that both groups consumed more on the devalued test days (moderate: 15.32 vs. 12.40 g, $F_{(1,53)} = 8.80$, MSE = 10.63, $\Delta = 7.47$, $P < 0.05$; extensive: 14.92 vs. 11.51 g, $F_{(1,53)} = 14.80$, MSE = 10.63, $\Delta = 13.24$, $P < 0.05$), but groups did not differ in overall consumption ($F_{(1,53)} = 0.41$, $P > 0.05$). If higher rates of consumption on devalued

test days caused greater general satiety, then rats should have consumed less during the preference tests on the devalued test days. This was true of the moderately trained group (10.10 vs. 7.15 g, $F_{(1,53)} = 16.66$, MSE = 7.30, $\Delta = 15.03$, $P < 0.05$), but not the extensively trained group (9.14 vs. 7.89 g, $F_{(1,53)} = 2.90$, $P > 0.05$). If differences in general satiety accounted for goal-directed responding, then the size of the difference in consumption between valued and devalued test days should positively correlate with the size of the LR sequence devaluation effect. The correlations for both groups were nonsignificant (moderate: $r = 0.16$, $P > 0.05$; extensive: $r = 0.27$, $P > 0.05$). Therefore, we do not think that the different levels of intake can account for the selective devaluation effects. More likely, they reflect the fact that rats are neophobic to relatively novel foods. Finally, groups did not differ in their percent preference for the nonsated pellet type during the preference tests (80% vs. 83% for moderate and extensive, respectively; $t_{(53)} = 0.57$, $P > 0.05$), indicating that the devaluation treatment was selective.

We found that moderately and extensively trained rats were reliably goal-directed, performing fewer target sequences when the outcome was devalued. This is despite the fact that by the end of training extensively trained rats performed with greater accuracy and were less variable in their sequence distributions and initiation times—consistent with them being more automatized. We also found that moderately trained rats were slower to initiate a sequence when rewards were devalued but did not show a change in the time to complete a sequence, while extensively trained rats showed the opposite pattern of behavior. This implies that goal-

directed control shifted from initiation to completion over training. According to one model, (Dezfouli and Balleine 2013), the decision to initiate a well-learned action sequence is thought to be controlled by a goal-directed process while the execution of the component parts is automatized and, thus, habitual. While our data confirm that early and late actions within a sequence are controlled by distinct decision-making processes (see also Morgan 1974; Balleine et al. 1995; Corbit and Balleine 2003; Balleine et al. 2005), our data are partly inconsistent with this particular model. Specifically, the finding that extensive training confers goal-directed control of sequence completion but not initiation seems problematic.

In summary, our data suggest that the general notion that automaticity leads to habit formation is overly simplistic. Action sequences can become automatized with overtraining, but goal-directed control remains and apparently shifts from its initiation to its completion. A representation of the outcome may become more restricted to the completion of the sequence with overtraining. We conclude that when a sequence becomes automatized that does not preclude it from being goal-directed, and based on the present set of data, it may not be sensible to equate habitual control with automaticity, or goal-directed control with a lack of automaticity.

References

Adams CD. 1982. Variations in the sensitivity of instrumental responding to reinforcer devaluation. *Q J Exp Psychol B* **34:** 77–98. doi:10.1080/14640748208400878

Balleine BW, Garner C, Gonzalez F, Dickinson A. 1995. Motivational control of heterogeneous instrumental chains. *J Exp Psychol Anim Behav Process* **21:** 203–217. doi:10.1037/0097-7403.21.3.203

Balleine BW, Paredes-Olay C, Dickinson A. 2005. Effects of outcome devaluation on the performance of a heterogeneous instrumental chain. *Int J Comp Psychol* **18:** 257–272. doi:10.1037/0097-7403.21.3.203

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. doi:10.1037/0097-7403.29.2.99

Desrochers TM, Amemori K, Graybiel AM. 2015. Habit learning by naive macaques is marked by response sharpening of striatal neurons representing the cost and outcome of acquired action sequences. *Neuron* **87:** 853–868. doi:10.1016/j.neuron.2015.07.019

Dezfouli A, Balleine BW. 2013. Actions, action sequences and habits: evidence that goal-directed and habitual action control are hierarchically organized. *PLoS Comput Biol* **9:** e1003364. doi:10.1371/journal.pcbi.1003364

Dickinson A, Nicholas DJ, Adams CD. 1983. The effect of the instrumental training contingency on susceptibility to reinforcer devaluation. *Q J Exp Psychol B* **35:** 35–51. doi:10.1080/14640748308400912

Jog MS, Kubota Y, Connolly CI, Hillegaart V, Graybiel AM. 1999. Building neural representations of habits. *Science* **286:** 1745–1749. doi:10.1126/science.286.5445.1745

Morgan MJ. 1974. Resistance to satiation. *Anim Behav* **22:** 449–466. doi:10.1016/S0003-3472(74)80044-8

Neuringer A. 2002. Operant variability: evidence, functions, and theory. *Psychon Bull Rev* **9:** 672–705. doi:10.3758/BF03196324

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. doi:10.1523/JNEUROSCI.1176-09.2009

Perlman MD, Rasmussen UA. 1975. Some remarks on estimating a noncentrality parameter. *Commun Stat* **4:** 455–468. doi:10.1080/03610927508827262

Rodger RS. 1974. Multiple contrasts, factors, error rate, and power. *Br J Math Stat Psychol* **27:** 179–198. doi:10.1111/j.2044-8317.1974.tb00539.x

Rodger RS. 1975. Setting rejection rate for contrasts selected post hoc when some nulls are false. *Br J Math Stat Psychol* **28:** 214–232. doi:10.1111/j.2044-8317.1975.tb00564.x

Rodger RS, Roberts M. 2013. Comparison of power for multiple comparison procedures. *J Methods Meas Soc Sci* **4:** 20–47. doi:10.2458/jmm.v4i1.17775

Rothwell PE, Hayton SJ, Sun GL, Fuccillo MV, Lim BK, Malenka RC. 2015. Input- and output-specific regulation of serial order performance by corticostriatal circuits. *Neuron* **88:** 345–356. doi:10.1016/j.neuron.2015.09.035

Yin HH. 2009. The role of the murine motor cortex in action duration and order. *Front Integr Neurosci* **3:** 3662–3669. doi:10.3389/neuro.07.023.2009

Yin HH. 2010. The sensorimotor striatum is necessary for serial order learning. *J Neurosci* **30:** 14719–14723. doi:10.1523/JNEUROSCI.3989-10.2010