Behavior chains reveal associative structure in human instrumental learning

Eric A. Thrailkill, Catherine R. Thorpe, and Mark E. Bouton
Department of Psychological Science
University of Vermont, Burlington VT, USA

•Contemporary theories suggest that instrumental (voluntary) behavior can be controlled by two processes: One goal-directed and sensitive to the value of the reinforcing outcome, and the other habitual and performed without the outcome "in mind" (e.g., Dickinson, 1985).

•Control by either process can be identified with the reinforcer devaluation method. After training the response, the value of the outcome is changed without the response, and then the response is tested without the outcome (extinction).

•Instrumental behavior is composed of distinct **consumption** and **procurement** behaviors.

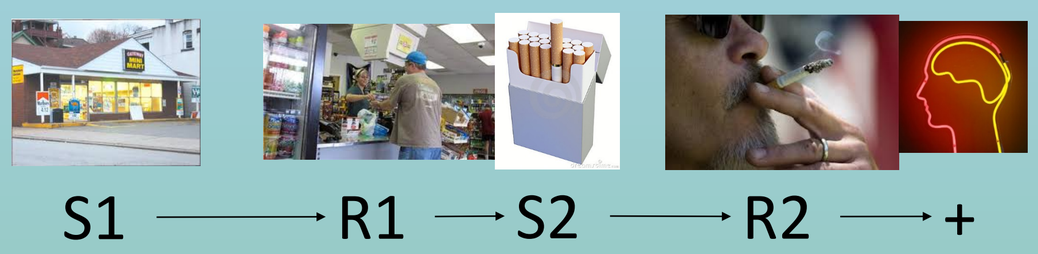
•For example, "cigarette smoking" refers to a behavior chain – one must buy cigarettes in order to smoke.

•Studies with rats suggest that chained responses form specific associations.

•In addition, the second response (R2) functions as the reinforcing "outcome" for the first behavior (R1) (Thrailkill & Bouton, 2016).

•Devaluation of a response with extinction can identify goal directed behavior, and some evidence suggests that additional training make the first response insensitive to extinction of the second response

•The present experiments 1) demonstrated that R1 can become a habit in a well-studied rat heterogeneous chain task, and 2) developed a computer-based task to study the associations learned and assess habit learning in human learning behavior chains.



Experiment 1
- 32 naïve female Wistar rats. ~90 days old.
- *Training.* R1 (e.g., lever press) in S1 (panel light near lever) turned on S2 (panel light near chain pull) which signals R2 (e.g., chain pull) will earn a reward (grain pellet).
- *Revaluation.* Half the rats received S2 trials in which making R2 could turn off S2 but did not earn food (extinction), and half were exposed to the chamber without trials.
- *Test.* All rats received S1 trials, making R1 could turn off S1 but S2 was not presented (extinction).

Results
- In Group Extended, extinction of R2 did influence R1 in the test.
- This suggests that extended training of a R1-R2 chain resulted in habit learning.
- In Group Brief, extinction of R2 resulted in reduced R1 in the test.
- This replicates previous results (Thrailkill & Bouton, 2016) and suggests that R2 is the "goal" for R1.

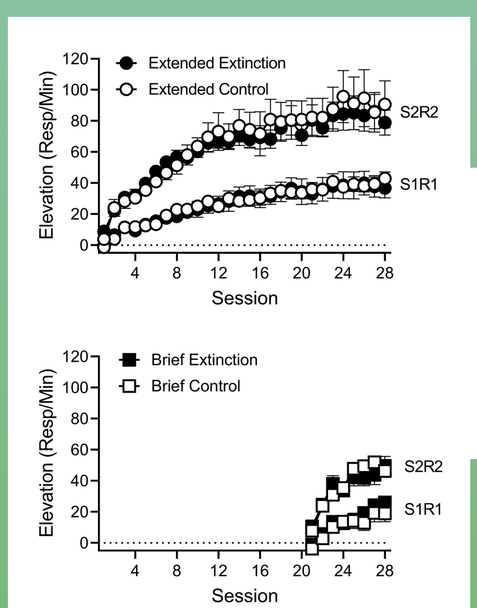
Experiment 1 assessed habit learning in our discriminated heterogeneous chain method with rats.

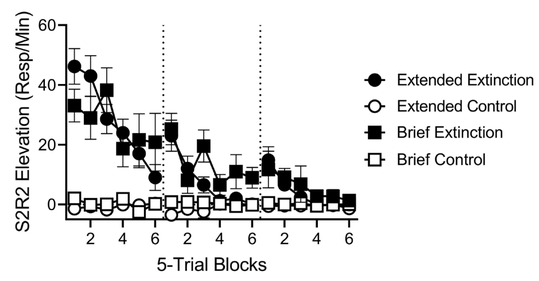
After brief training R1-R2 chain, extinction of R2 weakened R1. Consistent with habit learning, R1 was insensitive to R2 extinction after extended training (see also Zapata et al., 2010 for a related example).

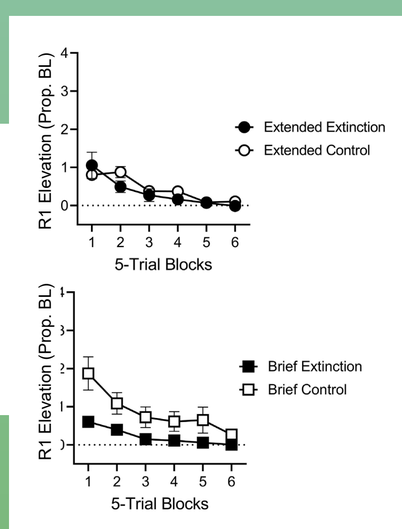
In rats, does extended chain training make R1 a habit?

Group	Training	Treatment	Test
Extended Extinction	S1R1-S2R2+ (28 days)	S2R2-	S1R1-
Extended Control		Chamber Exposure	
Brief Extinction	S1R1-S2R2+ (8 days)	S2R2-	
Brief Control		Chamber Exposure	

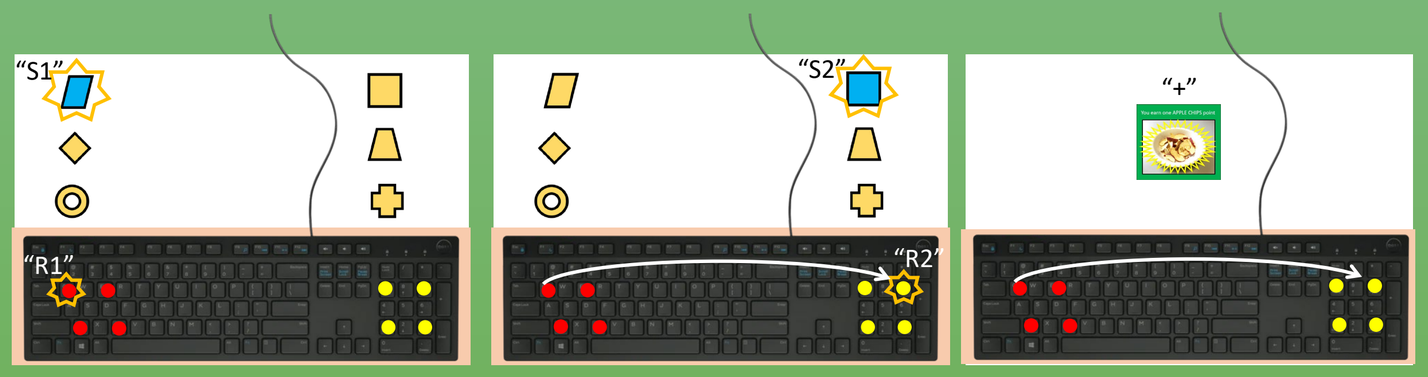
Note.
S = stimulus
R = response
+ = reinforced
- = nonreinforced (extinction)







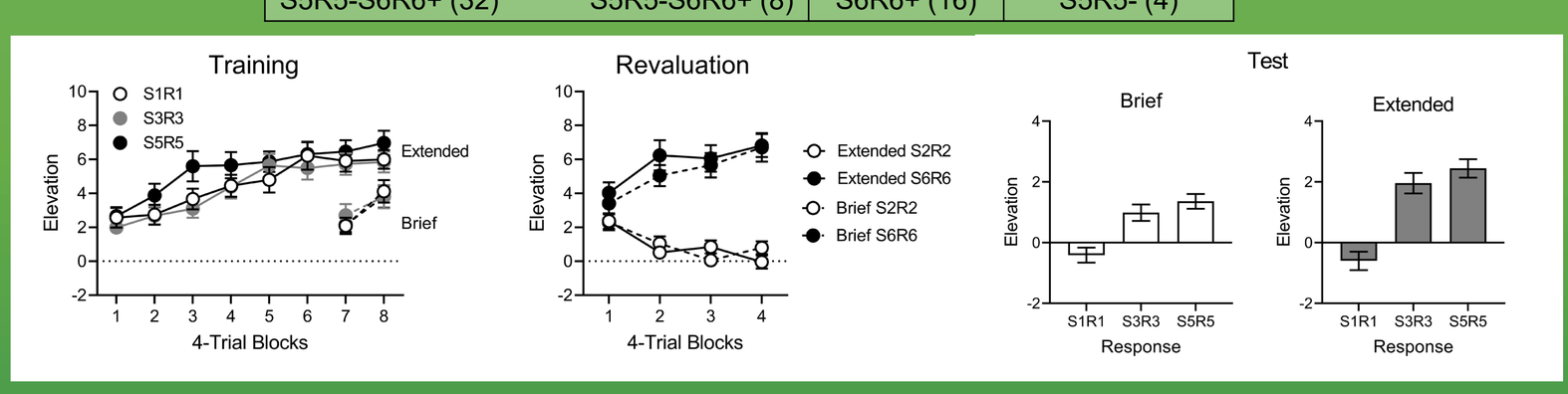
Experiment 2 (pre-registration: https://doi.org/10.17605/OSF.IO/S5FKB)
- 72 undergraduate students seeking course credit. 18-26 years old.
- *Training.* Computer version of heterogeneous chain task. Reinforcers were points associated with a preferred snack food item.
- *Revaluation.* Intermixed trials of S2 and S6. R2 was extinguished and R6 was reinforced on respective trials.
- *Test.* All received intermixed S1, S3, and S5 trials, making R1, R3, or R5 could turn off S but S2, S4, or S6 was not presented (extinction).



In humans, do chained responses form specific associations and does extended chain training make R1 a habit?

Training			Treatment	Test
Extended (*n*=35)	Or	Brief (*n*=37)		
S1R1-S2R2+ (32)		S1R1-S2R2+ (8)	S2R2- (16)	S1R1- (4)
S3R3-S4R4+ (32)		S3R3-S4R4+ (8)	-----	S3R3- (4)
S5R5-S6R6+ (32)		S5R5-S6R6+ (8)	S6R6+ (16)	S5R5- (4)

Results
- Group learned each chain during training.
- In revaluation, each group increased R6 (reinforced) and decreased R2 (extinguished).
- In the test, extinction of R2 weakened R1.
- R1 was *goal directed*
  - R1 was sensitive to the changed representation of its associated outcome, R2.
- Revaluation of R2 and R6 had a specific effect on the associated response.
  - R1 was selectively associated with R2, not R4 or R6.
- The revaluation effect on R1 remained after extended training.



Experiment 2 developed a human discriminated heterogeneous instrumental chain. After learning 3 two-link chains, extinction of R2 weakened its associated R1. This result replicates and extends findings with rats (Olmstead et al., 2001; Thrailkill & Bouton, 2016a).

Discussion
- **Rats and humans learn specific associations between responses in discriminated behavior chains** (Thrailkill & Bouton, 2016).
- Habit learning was observed after extended chain training in rats but not in humans. Habit learning is difficult to study in humans in instrumental tasks (de Wit et al.,2018).
- The experimental designs may be responsible; sensitivity to devaluation was compared between groups in Experiment 1 and within subject in Experiment 2.
- Concurrent or intermixed training prevents habit learning in rats (e.g., Colwill & Rescorla, 1985; Kosaki & Dickinson, 2010) and humans possibly (de Wit et al., 2018). Colwill and Rescorla (1985) found no evidence of habit learning after extended training with intermixed sessions of different response-outcome (R-O) combinations (R1-O1, R2-O2).
- Overall, the methods developed here identify a new approach to studying the common structure of instrumental behavior in humans and rats.

Contact Eric Thrailkill for reference list
Eric.Thrailkill@uvm.edu
@EricThrailkill posts learning articles on Twitter

This work was supported by NIH Grants K01 DA044456 to E. A. T. and R01 DA033123 to M. E. B.