Irrelevant-incentive learning and two-process theory

DOUGLAS S. GRANT, SHEILA M. GREER, and DONALD D. SEVERANCE
University of Alberta, Edmonton, Alberta T6G 2E9, Canada

Four groups of rats received one food (F) trial and one water (W) trial daily in a runway apparatus. Two of the groups received the incentives in the sequence FW, while the remaining two groups received the sequence WF. One group receiving each sequence was hungry in Phase 1, and thirsty in Phase 2; the remaining two groups were thirsty in Phase 1 and hungry in Phase 2. All four groups demonstrated negative transfer in Phase 2 relative to naive controls. This finding indicated that the animals failed to form expectancies involving specific rewards in Phase 1. It was concluded that (1) Trapold and Overmier's (1972) two-process theory of instrumental performance must be elaborated so as to specify the conditions under which animals do and do not form specific reward expectancies, and (2) the present procedure affords a technique whereby the nature of such conditions may be empirically determined.

Numerous theorists have proposed that an adequate theoretical treatment of learning phenomena requires the postulation of two associative processes (see Rescorla & Solomon, 1967, for a brief account of the historical development of such conceptualizations). These two processes may be characterized as the S-R associative process and the S-S associative process. Contemporary theoretical formulations within the two-process framework have emphasized the interaction between the two associative processes in controlling instrumental performance. However, two of the more prominent among these formulations differ in terms of the precise mechanism through which such joint control is exerted. According to Rescorla and Solomon (1967), Pavlovian S-S contingencies exert control over instrumental behavior through the arousal of a motivational state. On the other hand, Trapold and Overmier (1972) have argued that the mechanism of control involves the arousal of specific reward expectancies conditioned on the basis of S-S contingencies.

In support of Trapold and Overmier's (1972) proposition, Capaldi and Hovancik (1975) obtained evidence that rats learn expectancies involving specific rewards in an irrelevant-incentive learning task. Rats were run in a runway to an alternating sequence of two food (F) and two water (W) trials (FWFW) daily under thirst in Phase 2. The critical comparison involved the rate of development of appropriate Phase 2 patterning (i.e., faster on water trials than on food trials) as a function of Phase 1 treatment. Animals run to the same alternating sequence (FWFW) under hunger in Phase 1 developed appropriate Phase 2 patterning faster than those

placed on food deprivation but not given alley training. Capaldi and Hovancik concluded that the Phase 2 positive transfer resulted from the learning of expectancies involving specific rewards (i.e., food reward follows a water trial and water reward follows a food trial) during Phase 1.

Having found that animals can learn specific reward expectancies, Capaldi and her associates tested the assertion of Trapold and Overmier (1972) that such expectancies exert control in an instrumental training task (Capaldi, Hovancik, & Friedman, 1976). A standard transfer-of-control procedure was used in which two groups of rats received initial free operant barpress training for food reward: Group 1 received one pellet for each response and Group 5 received five pellets. During the subsequent noncontingent phase, animals received pairings of S1 with one pellet and S5 with five pellets in the absence of the response lever. Following the noncontingent pairings, training began on a discrete-trials procedure. Trials were initiated by the insertion of the response lever and terminated following a single response that also produced reward, again, one pellet for Group 1 and five for Group 5. In the final phase, insertion of the response lever was accompanied by S1 on some trials and by S5 on others.

Capaldi et al. (1976) reasoned that if S-S contingencies control instrumental performance through the arousal of a motivational state, responding to S5 should be faster than responding to S1 within both groups. That is, S5 should arouse a stronger motivational state associated with food, which would then summate with that aroused by insertion of the bar, leading to greater total motivation in the presence of S5. On the other hand, if the mechanism of control involves the conditioning of specific reward expectancies to the instrumental reaction, Group 1 should respond faster to S1, while Group 5 should respond faster to S5. That

is, during barpress training, the expectancy of one pellet is conditioned to responding in Group 1, whereas the expectancy of five pellets is conditioned to responding in Group 5. During the test phase, presentation of S1 in Group 1 and S5 in Group 5 arouses the expectancy conditioned to the reaction, and responding is facilitated. On the other hand, presentation of S5 in Group 1 and S1 in Group 5 arouses an expectancy not conditioned to the reaction, and performance is inhibited due to generalization decrement. The data supported clearly the latter predictions based on the notion that animals learn specific reward expectancies and that these expectancies can influence instrumental responding.

In an effort to assess the generality of Capaldi and Hovancik's (1975) findings, and hence the generality of Trapold and Overmier's (1972) formulation as well, Grant, Greer, and Severance (1979) attempted to determine whether the nature of drive would influence the formation of specific reward expectancies. To answer this question, Capaldi and Hovancik's procedure was employed with the addition of groups that experienced the different motivational states in reverse sequence, thirst in Phase 1 and hunger in Phase 2. The data of groups initially run hungry and shifted to thirst supported Capaldi and Hovancik's conclusion that hungry animals learn specific reward expectancies. However, the data of groups initially run thirsty and shifted to hunger failed to reveal any evidence that specific reward expectancies had been formed. This finding suggests that Trapold and Overmier's two-process theory may lack generality across qualitatively different drive levels within this procedure.

To examine further the boundary conditions under which rats form specific reward expectancies in an irrelevant-incentive learning task, the present experiment replicated that of Grant et al. (1979) with two important modifications. The first modification was suggested by an unpublished analysis of our prior data in which we examined separately performance on the first two daily trials and on the last two daily trials. The analysis revealed that patterning and transfer effects were considerably stronger on daily Trials 1 and 2 than on Trials 3 and 4. In fact, inclusion of the final two daily trials appeared to do nothing more than dampen the patterning and transfer effects. Based upon this analysis, a daily sequence of two, rather than four, trials was employed. It was anticipated that the use of the 2-trial/day procedure would lead to more pronounced transfer effects in drive-shifted groups. A second modification involved the manipulation of the order in which incentives were encountered within a day. Half of the hunger-thirst group and half of the thirst-hunger group received the daily sequence one food trial followed by a water trial (FW), while the remaining half of each group received the reverse sequence, WF. Our prior work revealed a strong positive transfer effect

for animals run hungry and shifted to thirst. For these animals the relevant incentive was encountered first and the irrelevant second on each day of Phase 1. On the other hand, animals run thirsty and shifted to hunger revealed strong negative transfer in Phase 2. Note that for these animals the order in which the relevant and irrelevant incentives were encountered within a day was the opposite of that for animals demonstrating positive transfer, irrelevant first and relevant second. It was hypothesized, therefore, that ordering of incentives might interact with drive to affect the formation of specific reward expectancies. If so, we might anticipate positive transfer when the hunger-thirst group received the FW sequence and when the thirst-hunger group received the WF sequence. Hunger-thirst animals run to WF and thirst-hunger animals run to FW would be expected to demonstrate negative transfer in Phase 2.

METHOD

Subjects
The subjects were 60 naive male albino rats of the Sprague-Dawley strain obtained from Holtzman Company. The subjects were approximately 90 days old upon arrival in the laboratory.

Apparatus
The apparatus was a wooden runway painted flat gray and measuring 200 x 10 cm with sides 23 cm high. On each trial the rat was placed in the startbox with the startbox door lowered and the goalbox door raised. When the rat oriented toward the door, the door was raised and a .01-sec clock was started. The clock was stopped when the rat interrupted a photobeam 150 cm beyond the startbox door.

Procedure
Upon arrival the animals were randomly assigned to one of six groups and received ad-lib food and water along with daily weighing and handling. Following this 7-day period, half of the animals were placed on a 23-h water-deprivation schedule (thirst, T), and the remaining animals were reduced to 85% of their ad-lib weights (hunger, H). During this second 7-day period, all animals continued to receive daily home cage experience with both food pellets and tap water placed in cups identical to those used in the goalbox of the runway.

The six groups may be referenced by the deprivation schedule in effect (H or T) and the training received (A, exposure to the alternating incentive sequence, or 0, handled but not given runway training) in Phases 1 and 2 (Phase 1 to the left of the hyphen and Phase 2 to the right of the hyphen). The six groups are thus designated HA-TA, TA-TA, HO-TA, TA-HA, HA-HA, TO-HA. Five animals in each group received the sequence one food trial followed by one water trial (FW), and the remaining animals received the sequence one water trial followed by one food trial (WF). Animals received 30-sec access to 3.0 cc of tap water or to 22 45-mg food pellets on water and food trials, respectively. Phases 1 and 2 were each in effect for 30 days, and 7 days intervened between phases to allow drive-shifted subjects to adjust to the change in deprivation schedule.

Dependent Measure
On each day a discrimination ratio (DR) was calculated for each animal according to the formula: $A/A + B$, where A is the run time on a day's food trial and B is the run time on a day's water trial. A DR of .50, therefore, reveals no pattern

running, while DRs above .50 indicate faster running on water trials and DRs below .50 indicate faster running on food trials.

RESULTS

Phase 2 discrimination ratios are shown in Figure 1. For groups run in Phase 1, performance on the last block of that phase is shown for comparison. The left panel presents ratios for animals run to the FW sequence, and the right panel presents ratios for animals run to WF. HA-HA and TA-TA were asymptotic at the end of Phase 1 and showed no systematic change in performance across blocks in Phase 2. Separate sequence by blocks ANOVAs failed to reveal any significant effects. The drive-shifted groups, HA-TA and TA-HA, demonstrated little or no change in patterning from the last block of Phase 1 to the first block of Phase 2, in spite of the change in drive. Patterning appropriate to the Phase 2 drive developed across blocks in each group. In HA-TA, appropriate patterning developed more quickly and reached a higher asymptote when the sequence was FW. In contrast, a slight tendency for the opposite effect was found in TA-HA: more rapid development of appropriate patterning and a higher asymptote when the sequence was WF. A sequence by blocks ANOVA performed on the data of HA-TA revealed a significant effect of blocks [F(9,72) = 30.49, p < .001] and a significant interaction [F(9,72) = 2.80, p < .01]. A similar analysis performed on the data of TA-HA revealed the main effect of blocks [F(9,72) = 36.24, p < .001] to be the only significant term. Finally, the naive groups, HO-TA and TO-HA, obtained ratios approximating .50 in Block 1 and gradually developed patterning across succeeding blocks. Both groups tended to develop patterning more quickly and to reach a higher asymptote under the FW sequence. A sequence by blocks analysis performed on the data of HO-TA revealed the main effect of blocks [F(9,72) = 16.32,

p < .001] to be the only significant term. A similar analysis performed on the data of TO-HA revealed significant effects of sequence [F(1,8) = 11.27, p < .01] and blocks [F(9,72) = 11.62, p < .001]. The interaction was also significant [F(9,72) = 3.51, p < .01].

Comparison of HA-TA and HO-TA reveals that naive animals were superior to drive-shifted animals throughout Phase 2, the difference being larger under the WF sequence. A groups by sequence by blocks ANOVA revealed significant main effects of groups [F(1,16) = 11.44, p < .01] and blocks [F(9,144) = 43.78, p < .001]. Three significant interactions were also found: Sequence by Blocks [F(9,144) = 2.18, p < .05], Groups by Blocks [F(9,144) = 3.15, p < .01], and the triple interaction [F(9,144) = 2.46, p < .05].

Comparison of TA-HA and TO-HA reveals that naive animals were generally superior to drive-shifted animals, more so under the FW sequence, except during the last four blocks, when the sequence WF was employed. The latter finding was produced by the fact that TO-HA developed patterning more quickly and reached a higher asymptote when the sequence was FW, due primarily to one animal that failed to pattern. A groups by sequence by blocks ANOVA revealed the same significant terms as the analysis performed comparing HA-TA and HO-TA: groups [F(1,16) = 14.38, p < .01], blocks [F(9,144) = 42.95, p < .001], Sequence by Blocks [F(9,144) = 6.07, p < .05], Groups by Blocks [F(9,144) = 5.46, p < .001], and the triple interaction [F(9,144) = 2.47, p < .05].

DISCUSSION

Regardless of whether animals were run to the sequence FW or WF, naive animals were consistently superior to drive-shifted animals throughout Phase 2. Therefore, animals shifted from hunger to thirst or from thirst to hunger failed to reveal evidence of the formation of specific reward expectancies. Of course, the failure to obtain positive transfer does not force the conclusion that specific reward expectancies were not formed during Phase 1, but only that, if formed, the facilitative effect of such expectancies on Phase 2 performance must have been masked by other factors. At this point, however, the most parsimonious account of the present findings is that specific reward expectancies were not formed.

The negative transfer demonstrated by HA-TA in the present experiment is curious in light of prior findings of positive transfer in hunger-thirst animals (Capaldi & Hovancik, 1975; Grant et al. 1979). The likely explanation for this anomaly is that a 2-trial/day procedure was employed in the present experiment, whereas a 4-trial/day procedure was used in studies demonstrating positive transfer in hunger-thirst subjects. Thus, while the inclusion of data from daily Trials 3 and 4 may only dampen patterning and transfer effects, the inclusion of such trials in the procedure appears critical to the formation of specific reward expectancies in HA-TA. These observations suggest that task complexity may determine, at least in part, whether or not specific reward expectancies are formed. In order to draw firm conclusions regarding the influence of task complexity on what is learned, further manipulation of this variable would be desirable. One possibility might involve employing an alter-
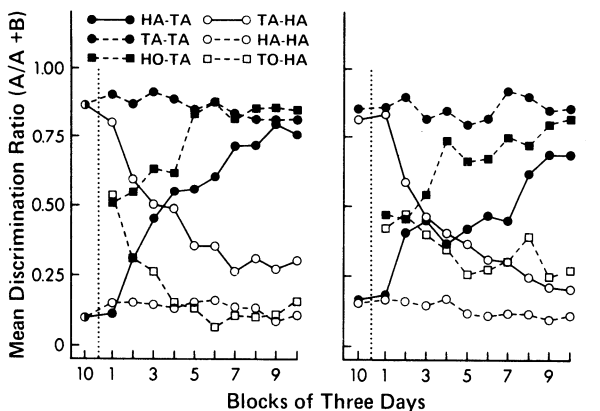

Figure 1. Mean group discrimination ratios in Phase 2. The left panel presents data for groups run to the FW sequence, and the right panel presents data for groups run to the WF sequence.

nating sequence and varying complexity by manipulating the number of trials given each day. A second possibility might involve giving all animals the same number of trials per day and manipulating complexity by exposing half of the animals to a single-alternation sequence and the other half to a double-alternation sequence.

The present findings and those of Grant et al. (1979) suggest strongly that the two-process analysis of instrumental performance offered by Trapold and Overmier (1972) may have limited applicability. That is, there appear to be situations in which specific reward expectancies are not formed and, therefore, in which instrumental behavior is controlled by the S-R associative process operating in isolation. On the other hand, the studies reviewed in the introduction demonstrate clearly that specific reward expectancies are formed in other situations (see also Capaldi, Smith, & Hovancik, 1977; Capaldi, Smith, & White, 1977). This suggests, then, that a set of rules or laws must be formulated that specify the conditions under which specific reward expectancies are formed. Factorial manipulation of type of drive, level of drive, ordering of incentives, desirability of incentives, and task complexity within the irrelevant-incentive learning procedure may well provide the empirical foundation for the postulation of such laws. Completion of this formidable research effort should contribute significantly to the development of a more adequate theoretical interpretation of instrumental behavior.

REFERENCES

CAPALDI, E. D., & HOVANCIK, J. R. Transfer from discrimination under hunger to discrimination under thirst: The role of expectancy and habit. *Learning and Motivation,* 1975, **6**, 230-240.

CAPALDI, E. D., HOVANCIK, J. R., & FRIEDMAN, F. Effects of expectancies of different reward magnitudes in transfer from noncontingent pairings to instrumental performance. *Learning and Motivation,* 1976, **7**, 197-210.

CAPALDI, E. D., SMITH, N. S., & HOVANCIK, J. R. Reversal learning as a function of changed reward location or changed drive. *Learning and Motivation,* 1977, **8**, 98-112.

CAPALDI, E. D., SMITH, N. S., & WHITE, L. A. Control of reward expectancies by drive stimuli. *Journal of Experimental Psychology: Animal Behavior Processes,* 1977, **3**, 178-188.

GRANT, D. S., GREER, S. M., & SEVERANCE, D. D. Irrelevant-incentive learning under hunger and thirst as assessed by discrimination transfer. *Canadian Journal of Psychology,* 1979, **33**, 51-56.

RESCORLA, R. A., & SOLOMON, R. L. Two-process learning theory: Relationships between Pavlovian conditioning and instrumental learning. *Psychological Review,* 1967, **74**, 151-182.

TRAPOLD, M. A., & OVERMIER, J. B. The second learning process in instrumental learning. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current theory and research.* New York: Appleton-Century-Crofts, 1972.