Behavioral Mixing and IRTs

In the section titled ***Behavioral processes in stimulus discrimination*** (***LCB***, *pp. 78 ff.*), we examined the behavior that occurs when a learner responds in an environment that is similar to, but differs from, the environment in which selection by reinforcement previously took place. A differential training procedure was used in which rapid responding was reinforced during one color and slow responding was reinforced during a second color. Rapid rates of responding were conditioned by reinforcing only those responses that occurred during a brief interval after the previous response; that is, responses with short interresponse times (*IRTs*). Slow rates of responding were conditioned by reinforcing only those responses that occurred after a longer interval since the previous response, that is, after long IRTs. Performance in a new environment was tested during an intermediate color (Collins, 1974). By establishing different rates during the two discriminative stimuli, the origins of responses that occurred during the intermediate test stimulus could be determined—that is, whether they arose from conditioning to the high-rate or to the low-rate stimulus. The finding was that responding during the intermediate test stimulus was at an intermediate average rate, but that this intermediate rate was the result of alternating bouts of rapid and slow responding (that is, short and long IRTs). The intermediate rate of responding was not produced by IRTs of intermediate duration. Thus, the new environment guided a *behavioral mixture* of previously conditioned responses, not new behavior.

The supplementary material presented here is directed at two issues: First, from a behavioral perspective, what does it mean to say that an IRT has been conditioned to a stimulus? Second, and more generally, how robust is the result that new environments evoke a mixture of "old" responses, not "new" responses?

The Status of IRTs as Operants

The technical definition of an IRT is unambiguous: The monitored response has occurred during some specified time (the criterion IRT) after the previous occurrence of the monitored response. For example, if a pecking response is reinforced on an IRT 3-s schedule, then the second of two successive key pecks must be separated by at least 3 s if the reinforcer is to be presented. Although the technical definition is unambiguous, its behavioral meaning is less so (Palmer & Donahoe, 1985). Does receiving a reinforcer after a given IRT guarantee that a temporally extended response, whose duration satisfies the IRT criterion, has been conditioned? Not necessarily.

Suppose that a response is occasionally reinforced after either of two IRTs—a shorter IRT and a longer IRT—and that no environmental stimulus is correlated with the two IRTs. Speaking casually, the learner has no way of "knowing" which IRT is scheduled for reinforcement. Suppose further that all of the possible reinforcers available during an experimental session are delivered to the organism. One possible interpretation of this finding is that two different temporally extended responses have been conditioned—one whose duration corresponds to the shorter IRT (e.g., two rapid discrete pecks) and one which corresponds to the longer IRT (e.g., a discrete peck followed by turning around in the chamber and then a second discrete peck). There is an alternative interpretation, however. Consider the alternative that the learner responds at a relatively high rate after a brief response, such as a key peck, has produced food. That is, food has functioned as a reinforcer for a discrete key peck. If the next food presentation were scheduled for the longer IRT, key pecks occurring close together in time would not produce the reinforcer and pecking would begin to extinguish. The decreasing rate of responding would eventually lead to a peck whose IRT met the longer IRT criterion. Key pecking would then

increase and, if a shorter IRT were next scheduled for reinforcement, food would be presented. However, if a longer IRT were next scheduled for reinforcement, responding would again extinguish and lead ultimately to a response that met the longer IRT criterion. This periodic conditioning and extinction of the discrete response of key pecking would allow the learner to collect all of the available reinforcers for both IRTs, but different temporally extended responses would not have been conditioned. Only discrete pecks would have been conditioned. Moreover, if conditioning and extinction occur more rapidly over repeated cycles—as has been demonstrated (Clark, 1972)—the discrete response would become more efficient in producing the reinforcer. It would appear that the learner had acquired two different temporally expended operants, but—in fact—only the acquisition and rapid extinction of a discrete response had occurred. No special appeal to temporally extended operants is required in such an analysis (Eisler, 1984).

The Robustness of Behavioral Mixing

Although interpreting the results of IRT procedures requires care, the technique does provide a powerful means to study temporally extended behavior. The technique is especially useful to understand behavioral processes in situations in which multiple stimuli are presented and those stimuli control different responses. The phenomenon of behavioral mixing during stimulus generalization illustrates this point. The prior reinforcement of different IRTs allowed us to see that the intermediate average rate of responding was the result of alternating between the different responses that had been conditioned to the training stimuli. The finding also illustrates the dangers inherent in averaging responses instead of examining the moment-to-moment changes in individual responses.

The conclusion that new environments engender a mixture of operants that were previously conditioned to similar environments has been reported many times, and does not depend exclusively on procedures involving the differential reinforcement of brief responses that terminate response sequences of different durations. (For reviews, see Donahoe & Wessells, 1980; Bickel & Etzel, 1985; see also ***LCB***, *footnote 1, p. 88.*) One procedure with an explicit temporally extended response used the duration of lever depression as an operant (Crowley, 1979). In the presence of a low intensity of light, reinforcers were occasionally presented if a rat held down the lever for a brief time. During a higher intensity of light, reinforcers occasionally occurred if the rat depressed the lever for a longer time. After the learner had acquired these discriminated *temporally extended* operants, testing was conducted in which lights of various intermediate intensities were occasionally presented. The durations of lever depressions were recorded during these tests.

Figure 1 shows the proportion of lever presses of different durations that occurred during the training stimuli (S1 and S5) and during the three intermediate test stimuli (S2, S3, and S4). During S1, long durations were reinforced; during S5 short durations were reinforced. The findings indicated that the relative frequency of different lever-press durations varied systematically as the test stimuli changed. For test stimuli closer in intensity to S1, longer durations predominated. For test stimuli closer in intensity to S5, shorter duration predominated. Thus the findings obtained with a procedure that required a temporally extended response also demonstrated behavioral mixing. Responding during intermediate environments was a mixture of previously conditioned responses, with the contribution of the various responses varying with the similarity of the test environment to the training environment.

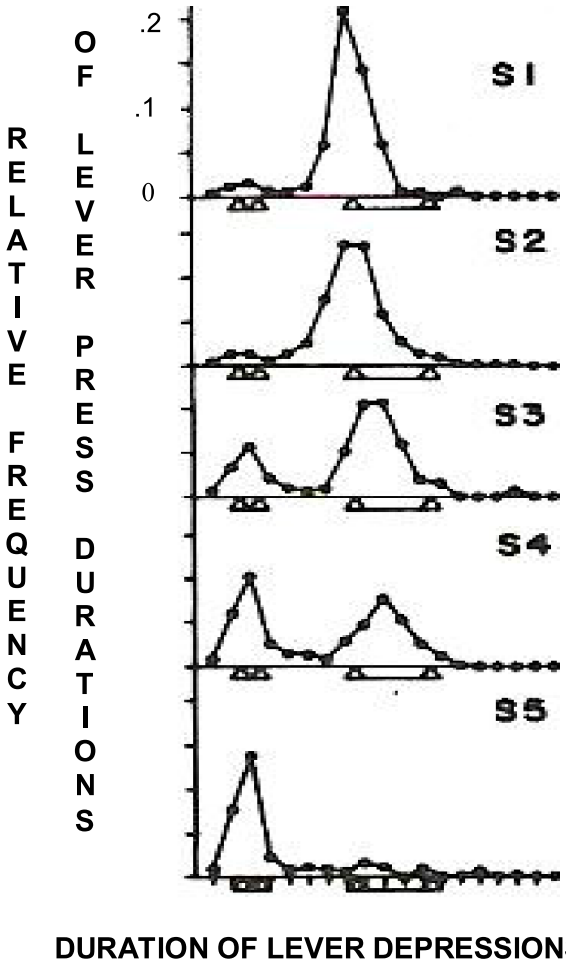

Figure 1. Changes in the frequency of durations of lever depressions during training stimuli (S1 and S5) and during intermediate test stimuli (S2, S3, and S4). (after Crowley, 1979)

The stimulus was a light that varied in intensity from bright (S1) to dim (S5). The pairs of triangles beneath the axes bracket the durations during which lever-press durations were reinforced. Longer depressions were reinforced during S1 and shorter depressions were reinforced during S5. Note that the proportions of total lever-press durations were greatest during the reinforced regions. No responses were reinforced during occasional presentations of test stimuli.

The finding that new environments evoke mixtures of previously conditioned behavior has important implications for complex behavior: The more varied the prior conditioning history, the richer is the mixture of environment-behavior relations upon which future reinforcers may act. An operant must occur before it can come into contact with new reinforcement contingencies. Learners with impoverished behavioral repertoires are genuinely handicapped; their behavior cannot profit from the "opportunities" (new contingencies) presented by new environments.

References

Bickel, W. K. & Etzel, B. C. (1985). The quantal nature of controlling stimulus-response relations as measure in tests of stimulus generalization. *Journal of the Experimental Analysis of Behavior*, **44**, 247-270.

Clark, F. C. (1972). Some procedural effects in operant discrimination reversal. *Psychological Record*, **22**, 83-101.

Collins, J. P. (1974). Generalization and decision theory. Unpublished doctoral dissertation. University of Massachusetts-Amherst, (described in J. W. Donahoe & M. C. Wessells, 1980, Learning, language, and memory. New York: Harper & Row)

Crowley, M. C. (1979). The allocation of time to temporally defined behaviors: responding during stimulus generalization. *Journal of the Experimental Analysis of Behavior*, **32**, 191-197.

Donahoe, J. W. & Wessells, M. G. (1980). Learning, language, and memory. New York: Harper & Row.

Eisler, H. (1984). Palmer, D. C. & Donahoe, J. W. (1985). Discriminated interresponse times: Role of autoshaped responses. *Journal of the Experimental Analysis of Behavior*, **44**, 301-313.