REPORT DOCUMENTATION PAGE WR-083 193 OMB No. 0704-0188

AD-A269 083

(2)

1. AGENCY USE ONLY (Leave

YPE AND DATES COVERED

4. TITLE AND SUBTITLE
Response acquisition [illegible]es:
A continuing quandary for molar models of operant behavior

5. FUNDING NUMBERS

6. AUTHOR(S)
Gregory Galbicka, Mary A. Kautz, and Traci Jagers

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Walter Reed Army Institute of Research
Washington, DC 20307-5100

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Development Command
Ft. Detrick, Frederick, MD 21703-5012

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

Accesion For
NTIS CRA&I ☒
DTIC TAB ☐
Unannounced ☐
Justification

By
Distribution /
Availability Codes

Dist	Avail and / or Special
A-1	20

DTIC ELECTE SEP 08 1993 S D E

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public release
Distribution Unlimited

12b. DISTRIBUTION CODE

DTIC [illegible] 1

13. ABSTRACT (Maximum 200 words)

The number of responses rats made in a "run" of consecutive left-lever preses, prior to a trial-ending right-lever press, was differentiated using a targeted percentile procedure. Under the nondifferential baseline, reinforcement was provided with a probability of .33 at the end of a trial, irrespective of the run on that trial. Most of the 30 subjects made short runs under these conditions, with the mean for the group around three. A targeted percentile schedule was next used to differentiate run length around the target value of 12. The current run was reinforced if i was nearer the target than 67% of those runs in the last 24 trials that were on the same side of the target as the current run. Programming reinforcement in this way held overall reinforcement probability per trial constant at .33 while providing reinforcement differentially with respect to runs more closely approximating the targe of 12. The mean run for the group under this procedure increased to approximately 10. Runs approaching the target length were acquired even though differentiated responding produced the same probability of reinforcement per trial, decreased the probability of reinforcement per response, did not increase overall reinforcement rate, and generally substantially reduced it.

14. SUBJECT TERMS
percentile schedules, molecular analyses, response differentiation, run length, response acquisition, response number, reinforcement probability, lever press, rats

15. NUMBER OF PAGES

16. PRICE CODE

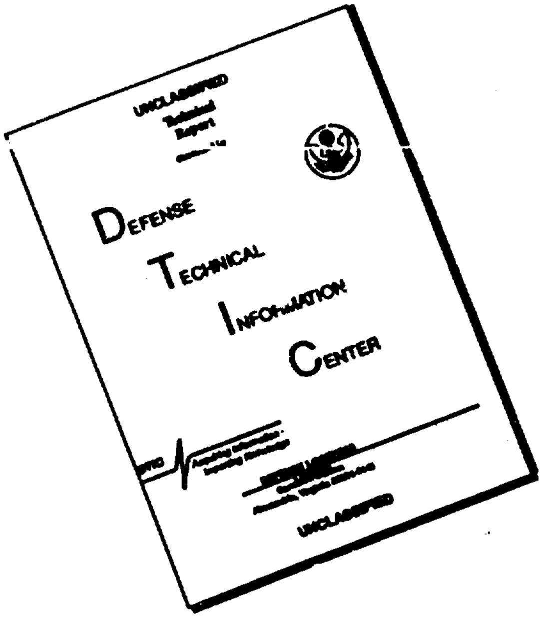
17. SECURITY CLASSIFICATION OF REPORT
18. SECURITY CLASSIFICATION OF THIS PAGE
19. SECURITY CLASSIFICATION OF ABSTRACT
20. LIMITATION OF ABSTRACT

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

ANNEX C

DISCLAIMER NOTICE



JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR 1993, 60, 171-184 NUMBER 1 (JULY)

RESPONSE ACQUISITION UNDER TARGETED PERCENTILE SCHEDULES: A CONTINUING QUANDARY FOR MOLAR MODELS OF OPERANT BEHAVIOR

GREGORY GALBICKA, MARY A. KAUTZ, AND TRACI JAGERS

WALTER REED ARMY INSTITUTE OF RESEARCH

15pf

The number of responses rats made in a "run" of consecutive left-lever presses, prior to a trial-ending right-lever press, was differentiated using a targeted percentile procedure. Under the nondifferential baseline, reinforcement was provided with a probability of .33 at the end of a trial, irrespective of the run on that trial. Most of the 30 subjects made short runs under these conditions, with the mean for the group around three. A targeted percentile schedule was next used to differentiate run length around the target value of 12. The current run was reinforced if it was nearer the target than 67% of those runs in the last 24 trials that were on the same side of the target as the current run. Programming reinforcement in this way held overall reinforcement probability per trial constant at .33 while providing reinforcement differentially with respect to runs more closely approximating the target of 12. The mean run for the group under this procedure increased to approximately 10. Runs approaching the target length were acquired even though differentiated responding produced the same probability of reinforcement per trial, decreased the probability of reinforcement per response, did not increase overall reinforcement rate, and generally substantially reduced it (i.e., in only a few instances did response rate increase sufficiently to compensate for the increase in the number of responses per trial). Models of behavior predicated solely on molar reinforcement contingencies all predict that runs should remain short throughout this experiment, because such runs promote both the most frequent reinforcement and the greatest reinforcement per press. To the contrary, 29 of 30 subjects emitted runs in the vicinity of the target, driving down reinforcement rate while greatly increasing the number of presses per pellet. These results illustrate the powerful effects of local reinforcement contingencies in changing behavior, and in doing so underscore a need for more dynamic quantitative formulations of operant behavior to supplement or supplant the currently prevalent static ones.

Key words: percentile schedules, molecular analyses, response differentiation, run length, response acquisition, response number, reinforcement probability, lever press, rats

Quantitative models of respondent (Pavlovian) conditioning have achieved a fair degree of success predicting trial-by-trial changes in responding (e.g., Rescorla & Wagner, 1972). Models of operant conditioning, on the other hand, have in general been silent with respect to response acquisition, concentrating instead on the order seen globally in response and time allocation of steady-state behavior as a function of relative reinforcement density (e.g., Davison & McCarthy, 1988). The analysis of operant acquisition is at somewhat of a comparative disadvantage, because those studying Pavlovian conditioning wield almost complete control over all experimentally relevant stimuli, but those studying operant conditioning traditionally surrender a degree of freedom to the subject by programming reinforcement contingent on behavior. As a result, the experimenter is incapable of precisely controlling the relation between behavior and environmental consequences, because the "free operant" is exactly that—free to vary from place to place, time to time, and subject to subject. This variation seemingly denies systematic analysis of the action of reinforcement at a local level. Skinner (1966), for example, noted that a learning curve "merely describes the rather crude overall effects of adventitious contingencies, and it often tells us more about the apparatus or procedure than about the organism" (p. 17).

Seven years after Skinner's (1966) pronouncement, John Platt developed the first in

---

The authors thank Timothy F. Elsmore, G. Jean Kant, and members of the Physiology and Behavior Branch for comments on an earlier version of this report. All research reported here was conducted in compliance with the Animal Welfare Act and other federal statutes and regulations relating to animals and experiments involving animals and adheres to principles stated in the *Guide for the Care and Use of Laboratory Animals*, NIH publication 85-23. All procedures were reviewed and approved by the WRAIR Animal Use Review Committee. The views of the authors do not purport to reflect the position of the Department of the Army or the Department of Defense (para 4-3, AR 360-5). Mary Kautz is now at the Division of Behavioral Biology, Hopkins Bayview Research Campus, 5510 Nathan Shock Dr. Suite 3000, Baltimore, Maryland 21224.

Request reprints from the first author, Department of Medical Neurosciences, WRAIR, Washington, DC 20307-5100 (or by e-mail: GAG@WRAIR-EMH1.ARMY.MIL).

a class of procedures (e.g., Alleman & Platt, 1973; Platt, 1973) that overcame the shortcomings noted by Skinner and allowed a systematic analysis of operant acquisition and differentiation. The *percentile* reinforcement schedules he devised make explicit the reinforcement contingencies involved in response shaping while simultaneously controlling either reinforcement probability or rate, holding one constant across the course of a differentiation within a single subject as well as across different subjects and response dimensions (e.g., Platt, 1984; see Galbicka, 1988, for a review). Because of the experimental control they afford, the constraints on the analysis of operant acquisition noted by Skinner (1966) are greatly attenuated, allowing an experimental analysis of how reinforcement effects response acquisition and differentiation.

The present study details some data from the differentiation of response number in rats under targeted percentile schedules. This arrangement controls the overall probability of reinforcement while differentiating response values around a fixed value, or target. The dimension of responding differentiated here was the number of presses made on the left lever of a two-lever operant conditioning chamber prior to a single press on the right lever. The left-lever pressing on each trial comprised a "run," and the percentile schedule differentially reinforced runs approximating a target of 12. This differential reinforcement was arranged by first determining whether the current run was shorter or longer than the target, and then comparing it to all prior runs within the most recent 24 trials that were likewise shorter (or longer, as the case may be) than the target. The reinforcement criterion was set such that two thirds of the comparison distribution fell outside the criterion zone, with the third closest to the target considered criterional (i.e., the criterional zone was above the 67th percentile of the distribution of runs shorter than the target and below the 33rd percentile of the distribution of runs longer than the target). This established a fixed probability of reinforcement equal to .33 at all times during the acquisition and maintenance of the differentiation for all subjects, independent of the absolute values of runs comprising the distribution at any particular time.

The present results demonstrate that reinforcement generates complex, tightly controlled behavioral sequences even when differentiated responding produces relatively little change in overall reinforcement probability, either leaves unchanged or reduces overall reinforcement rate, and increases the number of presses emitted per reinforcer. These effects hold true at all levels of meaningful aggregation—from entire conditions, to whole sessions, to blocks as short as 20 trials. As such, they illustrate that the relatively static quantitative formulations of operant behavior so far proposed, although very successfully describing some molar relations between aggregate behavior and reinforcement, can at best predict endpoints of more dynamic processes involving local reinforcement contingencies. Reinforcement changes behavior at a local level in such a way that subjects learn to emit complex patterns of behavior that decrease overall reinforcement density when doing so increases the *immediate* probability of food.

METHOD

Subjects

Subjects were 30 male Sprague-Dawley rats, fed freely to 350 g and maintained at that weight thereafter through restricted postsession feeding of chow. They were individually housed in acrylic rack-mounted cages lined with pine bedding, with freely available water in the home cage. The rack was removed from the colony room, which was maintained on a 12:12 hr light/dark cycle (onset time, 6:00 a.m.), at the same time every day and brought to the laboratory.

Apparatus

Sessions were conducted in five identically configured operant conditioning chambers (Coulbourn Instruments, Inc.). The instrument panel of each contained two response levers mounted symmetrically around an aperture (6.25 cm by 3.5 cm) in which reinforcers, consisting of a 45-mg food pellet (BioServe), could be delivered via a solenoid-operated pellet dispenser mounted behind the panel. The levers (Coulbourn Instruments Model E23-05 on the left and E21-03 on the right) required between 0.15 and 0.3 N to operate. No effort was made to standardize the force required across levers; however, each subject's box assignment remained constant, so the same requirement remained in force throughout the experiment. Each switch closure also operated a heavy-duty relay mounted behind the front wall above the food aperture. Above each lever were three lights (Sylvania 28ESB) mounted flush with the wall and covered with a red, green, or yellow cap. The floor of the chamber consisted of parallel stainless steel rods (0.5 cm diameter) spaced 1.8 cm, center to center. The chamber was entirely enclosed within a light- and sound-attenuating shell. White noise continuously present in the room helped further mask extraneous noise. A PDP® 11/73 minicomputer in an adjacent room, operating under the SKED11® (Snapper & Inglis, 1985) software system, programmed stimuli and collected data. The percentile schedule comparisons and calculations were evaluated by a set of FORTRAN subroutines (available upon request from the first author). Sessions were also monitored via Gerbrands (Model C-3SH) cumulative recorders.

Procedure

Following magazine training, during which pellets were delivered at random intervals averaging 30 s, pellets were delivered for any approach to and contact with either lever. Following this, pressing either lever produced a pellet. After 50 pellets, the procedure changed such that a green light was illuminated above one of the two levers, randomly selected on each trial, and only presses on that lever produced a pellet. This usually required a short period of remedial hand-shaping to move subjects from the preferred to the nonpreferred lever. After 100 presses under these contingencies, subjects moved rapidly between and pressed both levers. During the final pretraining condition, trials were signaled by illuminating the houselight and both green lights. A right-lever press following at least one left-lever press terminated a trial (right-lever presses prior to a left-lever press had no consequences) and initiated a 3-s blackout. Probability of pellet delivery following a trial was 1.0 during the first 33 trials, was .50 during the next 33 trials, and was subsequently reduced and maintained at .33 thereafter. This ultimate probability constituted the nondifferential reinforcement baseline and remained in effect for at least 15 sessions. During this and all subsequent conditions, sessions were conducted 5 days per week and lasted either 100 trials or 30 min, whichever occurred first.

The percentile procedure was then instituted, with a target value of 12 and a probability of a criterion run (w) of .33. Determining whether a run met criterion under this procedure involved three basic steps. First, the run was compared to the target to determine whether it was shorter or longer than the target. Next, the run was compared to all runs from the most recent 24 trials that were also short (or long, as the case may be) of the target. The number of such comparisons is denoted m. Finally, the run was considered criterional if it was closer to the target than k or the m comparison values, where $k = (m + 1)(1 - w) = .67\ (m + 1)$.

The mechanics of the above procedure involved initially determining the relative deviation of the current run from the target by subtracting the former from the latter. The first comparison value in memory (stored as a signed deviation from target, as well) was then multiplied by the current deviation to determine whether it was on the same side of the target (i.e., if the product was negative, the signs must be opposite, and that comparison was skipped). Deviations of zero (i.e., runs equal to the target) were arbitrarily classed as positive. If the deviations were both positive or both negative, the absolute values of the current and the comparison deviation were compared, and one of three counters was incremented, depending on whether the current deviation was closer to, equally distant, or further from the target than the comparison deviation. These steps were then repeated for each deviation in the comparison memory. This yielded tallies on each trial of the number of comparisons on the same side of the target with deviations larger, equal to, or smaller than the current one. The sum of these three tallies constituted the number of comparisons on the same side of the target (m) for that trial. The program first evaluated whether the current run was strictly closer than enough comparisons runs (the first tally) to exceed k, in which case it was considered criterional. Because the expression for k yields integer values only if $m + 1$ is a multiple of three, and the current deviation can only be closer to the target than an integer number of comparisons, k was rounded to the nearest integer. If the first tally did not exceed k, the number of equally distant deviations was added, and if this sum exceeded k, the run was considered criterional with a random probability equal to w (i.e., .33). Hence, even if all values in the memory equaled the present one, the present run would be considered criterional with a probability of .33. Independent of whether the current run was considered criterional, its signed deviation from the target replaced the oldest deviation in memory at the end of each trial (i.e., the memory always contained the most recent 24 deviations).

Because the conditional probability of reinforcement for criterional and noncriterional runs was 1.0 and 0.0, respectively, and criterional and noncriterional runs were mutually exclusive, criterional runs and reinforcement were isomorphic. Thus, not only did the overall probability of a criterional run remain controlled at the experimentally specified probability of $w = .33$ throughout acquisition and maintenance, so did the overall probability of reinforcement.

The number of deviations above or below the target in the comparison distribution varied across trials between 0 and 24. Allowing memory size to float is preferable to maintaining separate, fixed-sized memories for deviations above and below the target because the latter strategy can lead to comparisons to deviations no longer characteristic of present performance. That is, even if runs consistently deviated short of the target for hundreds of trials, the latter strategy would leave the memory for deviations above the target untouched, such that a run longer than the target would be evaluated with respect to this distribution even though it no longer accurately reflected performance.

Memory size affects the operation of percentile schedules in two ways. First, as memory size gets small, the estimation of percentiles suffers. That is, because m observations define $m + 1$ intervals into which the next run can fall, each observation represents the pth percentile of the distribution, where $p = 100/(m + 1)$. This places a lower limit on estimating criterional-response probability at $p/100$. Hence, for the percentile schedule to operate properly, a minimum number of comparison observations is necessary (here, to define the 33rd percentile, m must equal two or more). Second, memory size determines how long past behavior remains in the sample comprising the estimate of current behavior. As memory size increases, more remote runs contribute to this estimate. Occasional turnover in the comparison distribution is necessary to track any behavior change. Hence, memory size must be large enough to define necessary percentiles of the distribution accurately but small enough to allow frequent updating of the estimate of present performance. The memory size used here varied between trials from 0 to 24, allowing a maximum resolution of every 4th percentile while completely updating four times per session.

A final procedural variant was employed in an attempt to shape behavior symmetrically around the target. A symmetry routine like that described in Galbicka and Platt (1989, p. 151) was employed, in which the value of w was adjusted (w') depending on how much m differed from 12, the expected number of comparison values in a balanced memory. The routine is best understood by assuming a balanced memory and working backwards. If the comparison distribution was perfectly balanced, with 12 values above and below the target, then from the percentile equation $k = .67(13) = 8.71$, subsequently rounded to 9. Hence, any deviation closer to the target than the fourth smallest deviation would meet the criterion (i.e., would be closer than 9 other deviations). The symmetry routine, therefore, first classified *any* run as criterional if there were currently fewer than four comparisons on the same side of the target (i.e., if $m \leq 4$, $w' = 1.0$). As the comparison distribution size increased above 4, w was modified in direct proportion to the deviation from symmetry, such that $w' = 12w/m$ (i.e., for $4 \leq m \leq 12$, $1 \leq w' \leq w$; as the number of memory values approached symmetry, w' approached w). As memory size increased above 12 (i.e., the present run fell on the preferred side of the comparison distribution), fewer runs than nominally programmed were considered criterional (i.e., for $m > 12$, $w' < w$). This strategy becomes self-defeating, however, as comparison values overwhelmingly predominate on one side of the target (i.e., if $m = 24$, $w' = 1/2w$), as they would early in acquisition. This adjustment, therefore, was used only when the number of comparisons on the *nonpreferred* side exceeded 4 (and hence $4 \leq m \leq 19$). For m

> 19, the quantity $(1 - w)$ in the percentile equation was multiplied by $24/m$. At the point of transition between these two algorithms, both specify $w' = w/(2 - w) = .197$, but the latter specifies w' approaches w as m approaches 24, restoring criterional response (and reinforcement) probability to the expected value.

Under all conditions, the time of every stimulus event and every lever press was recorded such that the entire session could be reconstructed to the nearest 0.01 s. Data were subsequently transferred to a minicomputer (Digital Equipment Corporation) for storage and analysis.

RESULTS

Figure 1 shows overall mean run (left response per trial) for the group across sessions under the nondifferential baseline and targeted percentile conditions, as well as the mean run reinforced. The mean run under baseline was generally short (approximately three), and relatively stable. The mean reinforced run did not systematically differ from the overall mean, demonstrating the nondifferential nature of the baseline reinforcement contingency. Under the percentile schedule, mean run length increased rapidly, reaching an asymptotic level of just over 10 in approximately 20 sessions. Note that, as required by the percentile procedure, the mean reinforced run also increased steadily, remaining consistently closer to the target than the mean run overall.

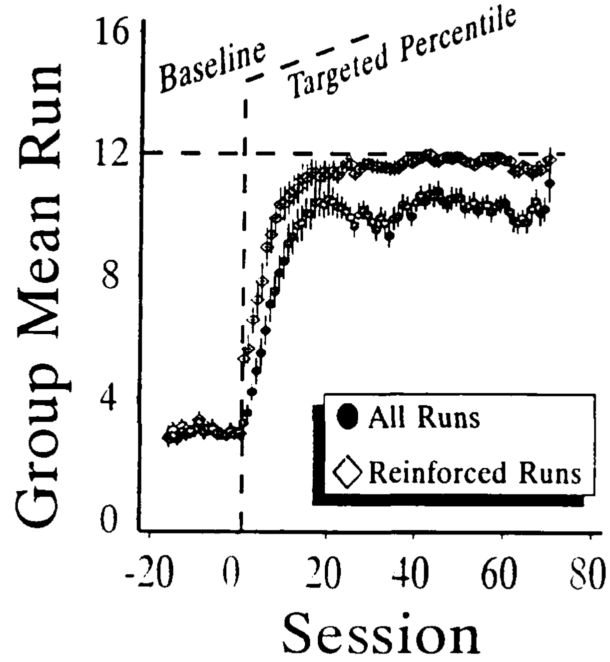


Fig. 1. Run length (left responses per trial) for all runs (closed circles) or reinforced runs only (diamonds) for the group across sessions. Points and vertical bars are means ± *SEM* of individual-subject session means. Values to the left of the vertical dashed line were obtained under the nondifferential reinforcement baseline, those to the right under the targeted percentile schedule. The dashed horizontal line represents the target during the latter.

To provide a gross measure of how this change in the group mean reflected changes in individual performance, Figure 2 presents the cumulative percentage of subjects attaining various acquisition criteria as a function of time under the percentile schedule. To derive these values, every session was first divided into five 20-trial blocks, and then the entire sequence was scanned for 25 or 50 consecutive blocks, during which the mean run for a particular subject remained at or above either 50%, 67%, or 75% of target. The block size was set at 20 trials to provide the minimal aggregate over which various other measures of behavior and reinforcement could evince a range of meaningful values (i.e., values that could potentially demonstrate substantial variability for reasons other than small sample size). The block in which the 25th (or 50th) consecutive block occurred constituted the acquisition block for that subject; hence, the minimum value was 25 (or 50). The fastest subject met the 50% and 67% criteria shortly after the minimum, irrespective of the number of consecutive blocks required, and met the 75% criterion for 25 consecutive blocks after just over 50 blocks (during the 11th session) and for 50 consecutive blocks just prior to the 100th block. All but 2 subjects met the 50% criterion for 25 consecutive blocks within 100 blocks, whereas 80% of the subjects met the 67% criterion and 40% met the strictest criterion for 25 consecutive blocks within the same period. After 50 sessions (250 blocks), just over 70% of the subjects had met the 75% acquisition criterion for 25 consecutive blocks. The required number of consecutive blocks interacted with the percentage of target required in determining the percentage of subjects meeting acquisition. The percentage of subjects attaining the 50% criterion was only slightly decreased by increasing the number of consecutive blocks required, with over 80% meeting the criterion for 50

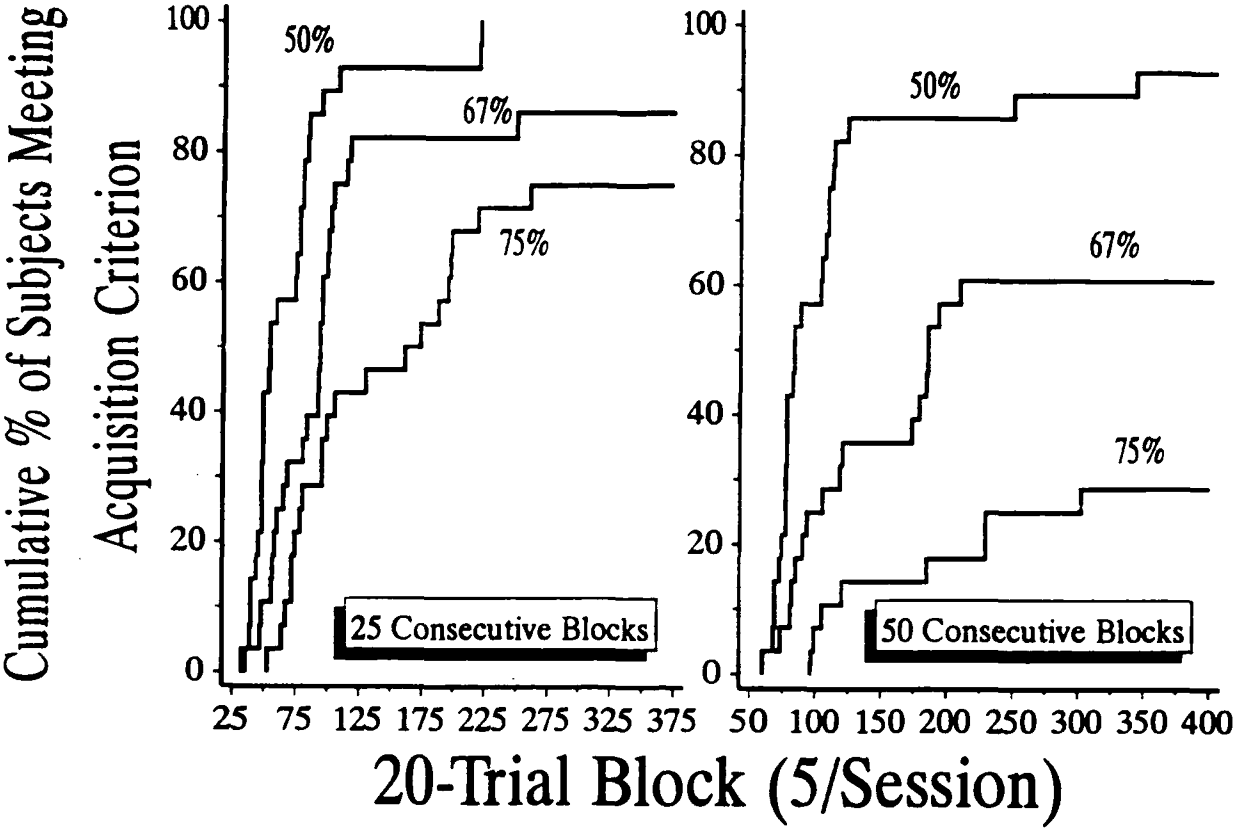


Fig. 2. Cumulative percentage of subjects maintaining a minimum mean run of 50%, 67%, or 75% of target for either 25 (left panel) or 50 (right panel) consecutive 20-trial blocks (five blocks per session) as a function of consecutive block number under the percentile schedule. The lines increment during the session in which the 25th (or 50th) block occurred.

consecutive blocks by the 100th block. Only 60% of the subjects maintained run lengths equal to or greater than 67% of the target for 50 consecutive trials within the first 200 blocks, compared with over 80% for the 25-block criterion, whereas the percentage of subjects meeting the 75% criterion for 50 consecutive blocks was reduced even more over its 25-block counterpart, with only 20% meeting criterion (compared to 70%) within the first 250 blocks.

Figure 3 shows mean run (overall and reinforced) across 20-trial blocks for each of 4 subjects, selected to illustrate characteristics of the percentile procedure as well as of responding. Subjects 38 and 39 showed fairly typical acquisition under the percentile procedure. Run length gradually increased to a value slightly lower than target, during which time the mean run reinforced increased as well to remain longer than the overall mean. As run length increased above the target, however, the mean reinforced run remained displaced nearer the target, such that it was now relatively shorter than the overall mean (e.g., Subject 39's data during Blocks 90 through 100). Run length subsequently decreased below the target, such that reinforced runs were now relatively longer than the mean, and the cycle repeated, with noticeable oscillation in run length. For Subject 38, these oscillations appeared as almost a sawtooth pattern, whereas for Subject 39 transitions were more gradual (the inset in each panel expands several cycles for each subject). Subject 40's results demonstrate that these oscillations did not always occur, and that not only did the mean reinforced run increase with increases in overall run length to the target value but it also decreased to track decreases in overall run length, both during the long sequence between Blocks 25 and 50 and during the single blocks at approximately Blocks 175 and 220, for example. In all these instances, however, the mean reinforced run always remained closer to the target than the mean run on that block, maintaining the differential reinforcement contin-

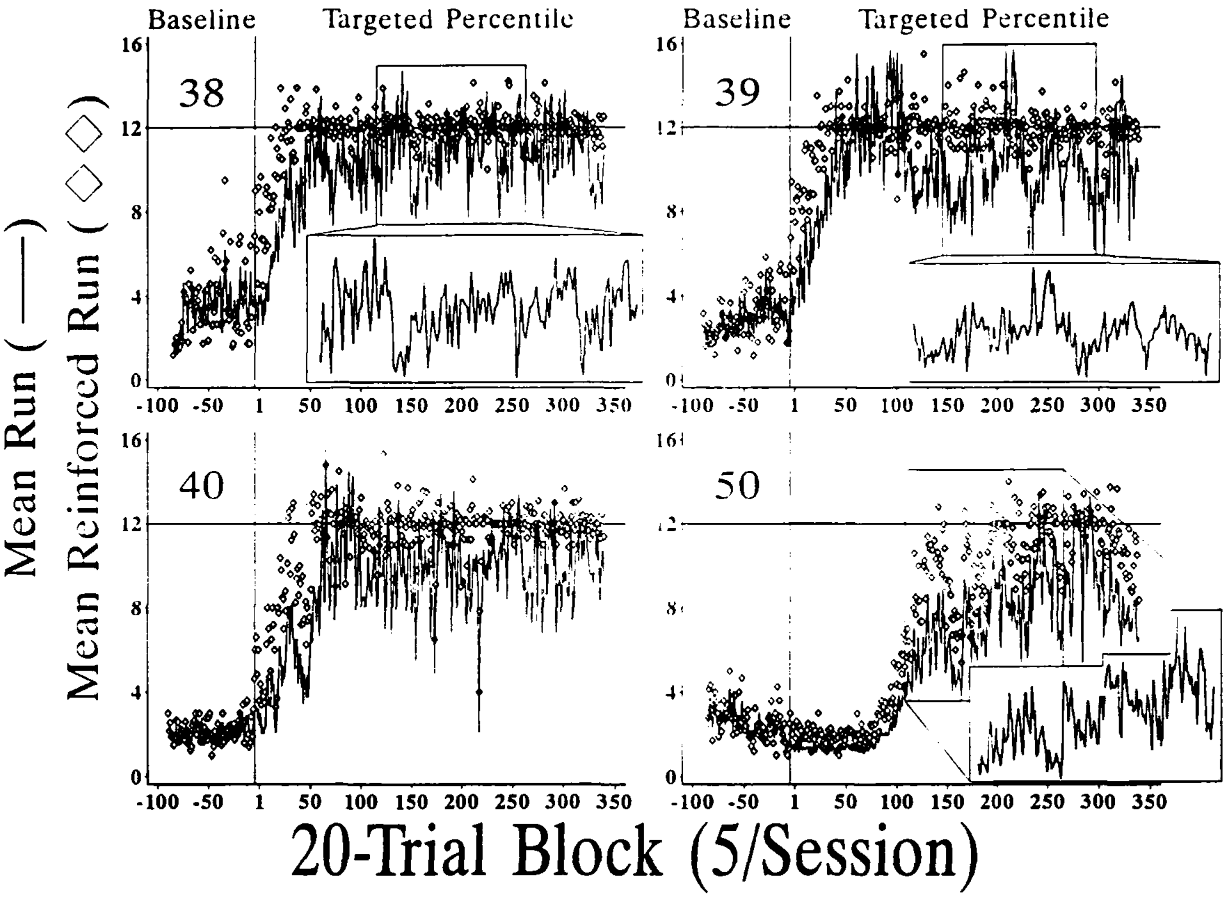


Fig. 3. Mean run (left responses per trial) on all trials (connected lines) or reinforced trials only (diamonds) for 4 subjects (separate panels) during consecutive 20-trial blocks of baseline (left of the vertical in each panel) or the percentile schedule (right of the vertical). The horizontal dashed line indicates the target during the percentile schedule. The insets in the panels for Subjects 38, 39, and 50 expand several cycles of run-length oscillation.

gency. Finally, Subject 50's data present an extreme example of delayed acquisition. Other than the extended period of near-invariant short runs for the first 75 blocks, however, there was little to distinguish this subject's data once acquisition began. It occurred more gradually than for Subjects 38 and 39, but this was also true of other subjects. Note that throughout the targeted percentile procedure, even before runs began to change appreciably for this subject, reinforcement remained differentially contingent on runs closer to the target, albeit by a slender margin.

One factor that might influence time to complete acquisition is the amount of variability present in the baseline run distribution from which the percentile schedule selects criterional runs. An inverse relation might be expected, such that less variability under baseline would correlate with more extended acquisition. This expectation was only partially borne out by the present data. Table 1 shows correlation coefficients (r) between the standard deviation of runs from the last five baseline sessions for each subject and the session on which that subject met each of the different acquisition criteria presented in Figure 1, further classified by whether acquisition occurred within 150 blocks or 400 blocks. Also shown are the probabilities by which each coefficient differed statistically from zero (p) and the number of subjects on which the correlation was based. A relatively strong inverse correlation was apparent between run variability and time to acquisition at both 50% criteria for subjects acquiring by the 150th block. Extending the window to the 400th block weakened both correlations, although the one for the 25-block criterion remained relatively substantial ($p < .05$). Correlations based on the 67% criterion were generally smaller than their 50% counterparts, except for those based on

Table 1

Pearson product moment correlations (r) between individual subjects' run-length standard deviations during the last 5 days of baseline and the block on which they met the six different acquisition criteria, along with the probability that the coefficient equaled zero (p) and the number of subjects on which each correlation was based (N). The rightmost columns present correlations obtained using all subjects that acquired the differentiation at the different levels by the 400th block, and the middle three columns are correlations based only on those subjects that achieved acquisition within the first 150 blocks.

Criterion		Subjects meeting criterion					
		By 150th block			By 400th block		
% Target	Block	r	p	N	r	p	N
50	25	−0.51	.01	26	−0.44	.02	28
50	50	−0.55	.01	26	−0.30	.13	26
67	25	−0.37	.08	23	−0.33	.11	24
67	50	−0.66	.04	10	0.08	.75	17
75	25	−0.32	.29	13	0.18	.44	21
75	50	0.17	.83	4	0.78	.02	8

subjects reaching the 50-block criterion by the 150th session; these did achieve statistical significance ($p < .05$). Correlations based on the 75% criterion were generally insignificant, except for the correlation based on subjects reaching the 50-block criterion within the larger window. This yielded the largest and only significant *positive* correlation coefficient of any condition ($r = 0.78, p < .05$). Hence, it appears that baseline variability may help predict an initial, relatively small change in the direction of the target, but not the time to fine-tune a differentiation around a particular target value. This interpretation, of course, should be tempered by the small sample sizes on which the significant 67% and 75% correlations were based.

To provide an indication of how different behavioral measures concurrently changed and to present data for some additional subjects, Figure 4 shows five different measures plotted across 20-trial blocks for 6 subjects (Subject 50's run-length data were also presented in Figure 3). The measures were chosen such that they could simultaneously be presented on semilogarithmic axes with minimal overlap. They are, in order of increasing frequency, reinforcement rate, reinforcement probability, response rate, run length, and trial rate. Subjects 34, 43, and 53 show the most typical acquisition pattern. Imposition of the targeted percentile procedure increased run length rapidly from a mean between two and three to a value that oscillated between eight and 14. Reinforcement probability remained relatively constant throughout this change in run length. This increase in presses per trial most often occurred concomitant with an increase in response rate, although for Subject 34 this rate increase was slightly delayed. The increased response rate, however, seldom compensated for the increase in the mean run, such that the rate of trial completion decreased drastically to around half its baseline value. Because reinforcement probability was experimentally controlled, this decrease in trial rate concomitantly decreased overall reinforcement rate. Subject 55 was one of the few subjects for whom response rate increased parallel to the increased number of responses per trial, keeping the rate of trial completion (and hence reinforcement) constant. Subject 50's results are again striking because of the delay in acquisition. Mean run length was decreasing for this subject during baseline, and imposing the targeted percentile schedule did not reverse this trend, most immediately resulting in almost complete minimal runs on each trial (i.e., runs of one). Response rate stabilized during this time such that the rate of trial completion approached 30 trials per minute, generating a high and stable reinforcement rate as well. After approximately 15 sessions, and despite the existing high rate of reinforcement, acquisition finally commenced, and although response rate increased substantially during this period, trial and reinforcement rates were driven down by almost two thirds as mean run approached the target.

Subject 56 was the only subject who failed to maintain differentiated runs in the vicinity of the target. As run length increased from around three to about 12 after 10 sessions under the percentile procedure, response rate, which was already relatively high (two responses per second), increased by only about one third. As a result, trial rate and reinforcement rate plummeted. During the next 15 sessions, run length decreased, increasing trial and reinforcement rates. This was followed by a subsequent increase in run length for approximately 10 sessions, with a correlated decrease in trial and reinforcement rates. Thereafter, run length consistently decreased to near

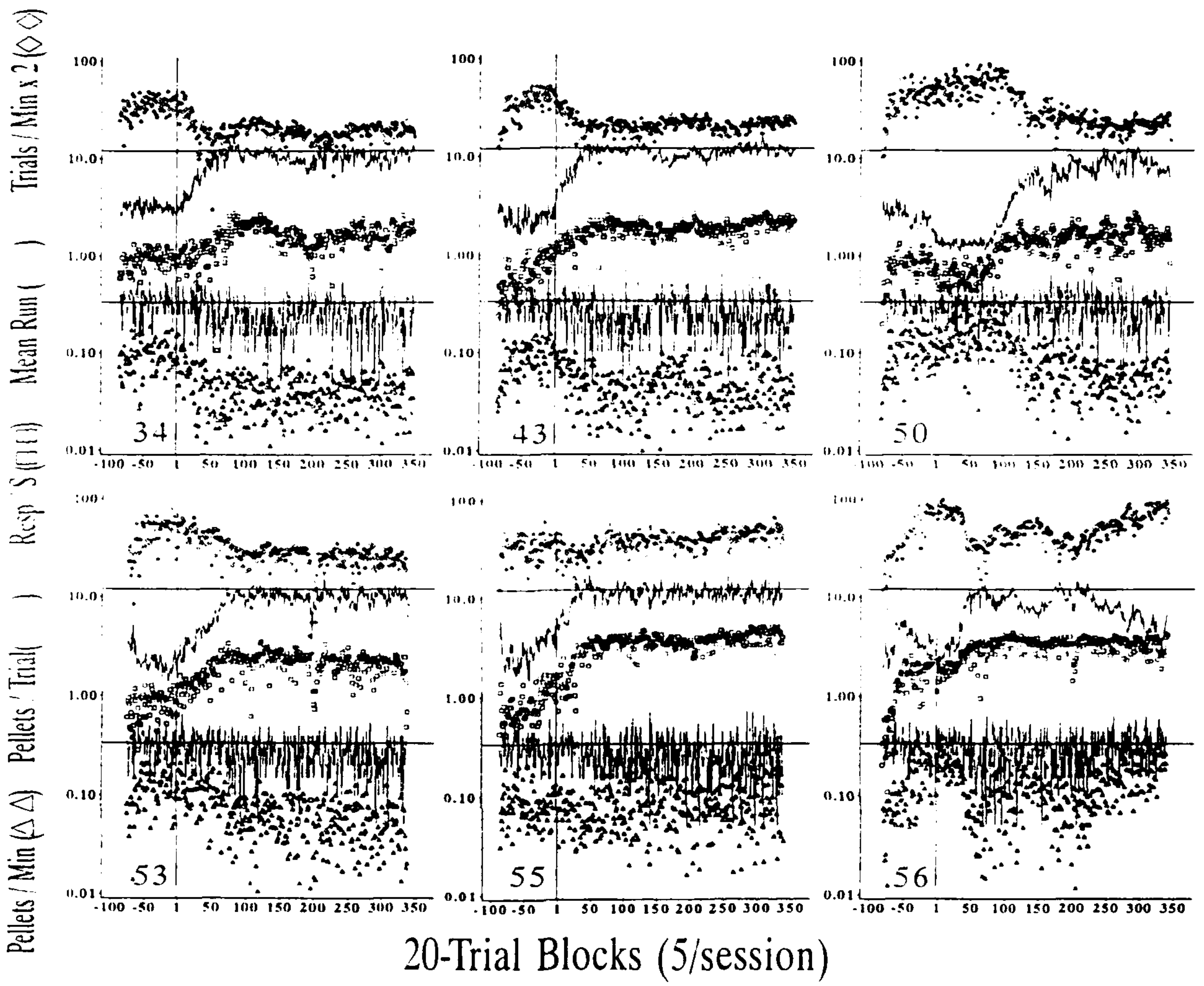


Fig. 4. Trial rate (trials per 2 min: diamonds), run length (left responses per trial: solid line), response rate (responses per second: squares), reinforcement probability (pellets per trial: dashed line), and reinforcement rate (pellets per minute: triangles) for each of 6 subjects (individual panels) under the baseline and percentile procedures (left and right of the vertical in each panel). Values represent block means. Note the semilogarithmic axes. Horizontal lines indicate the percentile target (upper line) and the expected reinforcement probability (lower line).

baseline values, restoring trial and reinforcement rates to the high values obtained prior to the short-lived differentiation.

A close look at the reinforcement probabilities in Figure 4 reveals a small but systematic decrease below the value programmed, correlated with periods when mean runs were slightly below the target. This decrease was evident for Subjects 34, 43, and 55 from approximately Block 50, and for Subject 53 from Block 75 onward, except for the period between Blocks 150 and 200 for Subject 43, during which mean runs fell even further below the target. For Subject 50, the decrease in reinforcement probability was not evident except for the short period between Block 275 and 300, during which the mean run remained very close to, but short of, the target. For Subject 56, variability in the mean run made detecting a consistent decrease in reinforcement probability difficult; however, after runs began to decrease consistently (approximately Block 225), reinforcement probability became less variable and showed no decrease. These variations from the nominal probability programmed by the percentile schedule likely resulted from the memory symmetry routine, which operated only after runs longer than the target comprised a portion of the comparison distribution. When all runs fell short of the target early in acquisition, the routine did not operate. Once runs above the target were oc-

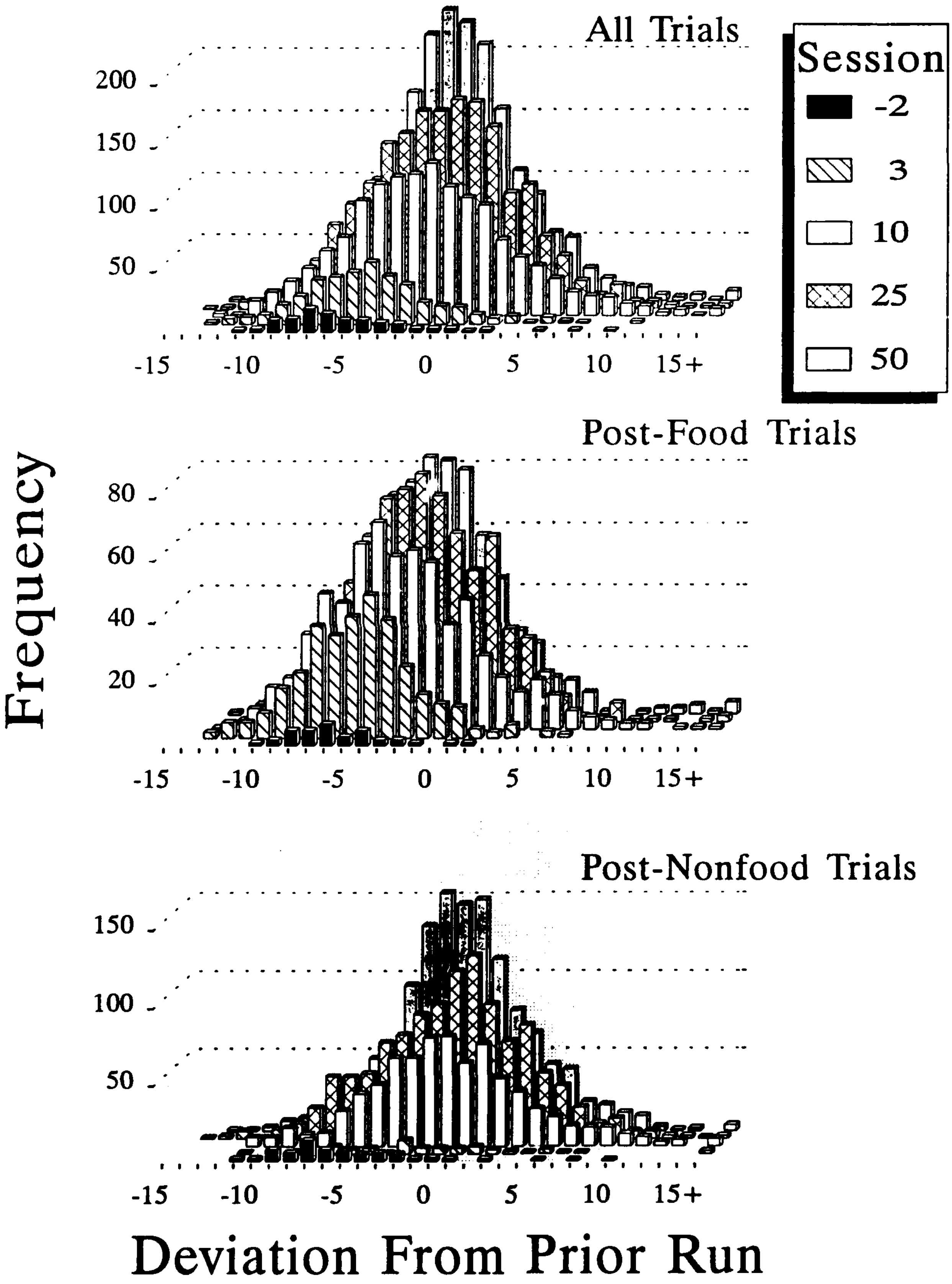


Fig. 5. Frequency distributions of deviations from the previous run, limited to those trials following a run of between 8 and 16, during the penultimate session under baseline (Session −2) and during Sessions 3, 10, 25, and 50

casionally emitted, however, criterional response probability was reduced for runs on the preferred side (below target) and incremented for runs on the nonpreferred side. If this restored the distribution to symmetry, the resulting probability of a criterional response would be the nominal value (w). However, because the distribution remained asymmetrically positioned below the target, most runs were selected with an adjusted probability $w' < w$, and reinforcement probability remained slightly reduced.

To examine local changes in runs at different points during differentiation, deviations between successive runs (i.e., the difference between the current and the previous run) were computed for every subject during the penultimate session under baseline (−2), and the 3rd, 10th, 25th, and 50th sessions under the percentile procedure. Because run length is bounded by a physical minimum and most likely by a behavioral maximum, deviations between successive values are likewise constrained (e.g., a distribution comprised solely of small runs cannot have large negative deviations). To minimize the effects of these constraints and provide a less biased measure, deviations were determined only if the run on the reference (preceding) trial was between eight and 16. The top panel of Figure 5 shows the frequency of all deviations for the group, and the bottom two panels segregate deviations by whether food was presented on the reference trial. Absolute, as opposed to relative, frequencies are presented to indicate changes in the number of observations comprising each distribution, as well as how those deviations were distributed. Given the differences in total observations between distributions, however, comparisons should emphasize relative shapes and not absolute frequencies. Under baseline and the third percentile session, most deviations were negative. This was not surprising because the minimum run on the previous trial was eight and the mean run at this time was around three (see Figure 1). As the differentiation progressed, the upper tail of the overall distribution extended to include more positive deviations. The mode ultimately settled at −1 and appeared relatively symmetric. Deviations following criterional runs between eight and 16 (middle panel) were shifted toward negative deviations. Conversely, distributions of deviations following noncriterional runs between eight and 16 had relatively larger numbers of positive deviations, with a mode of 0 and +1 during the 10th and 25th sessions and positively displaced secondary modes during the 10th, 25th, and 50th sessions.

DISCUSSION

All models of behavior that discount the influence of local reinforcement contingencies in deference to aggregate relations predict that runs should have remained short throughout this study, because such runs maximize trial rate and/or minimize the number of responses per reinforcer. Maximizing trial rate maximizes reinforcement rate, because reinforcement probability per trial was constant throughout. Minimizing responses per trial increases reinforcement probability per response or decreases the "price" of food (cf. Hursh, 1980). Each of these is easily accomplished by responding once on the left lever and then switching to the right lever to end the trial.

Of the 30 subjects studied, however, the behavior of only 1 even remotely approached this prediction. Most subjects made runs longer than one under both the baseline and the percentile procedures. No doubt, these models could be modified to allow the variability induced by intermittent reinforcement under both procedures to predict runs longer than the absolute minimum, but this cannot account for the differential results under the two procedures. Under the nondifferential baseline, when there was no local contingency with respect to run length, subjects approximated the minimum allowable run by making relatively short runs. But subjects in the present study overwhelmingly acquired differentiated responding when the targeted percentile procedure was instituted, making not merely longer runs but runs in the vicinity of an experimentally defined target, even though doing so did

---

←

under the targeted percentile procedure. The top panel presents all deviations, and the bottom panels segregate deviations depending on whether the previous trial ended in food. See figure legend for session identification. Values are total frequencies for the group.

not increase reinforcement *probability* (either per trial or per response), required *more* responses per pellet, and resulted either in the same reinforcement *rate* (at best) or often severely reduced it. Of the 30 subjects in this study, only 1 avoided being "trapped" by the percentile schedule into emitting a response pattern that did not optimize aggregate reinforcement parameters. Further, the present subjects represent only the most recent ones to be exposed to the contingencies described here. Runs of over 100 subjects have now been differentiated under targeted percentile schedules like the present one, with similar results (cf. Galbicka, Fowler, & Ritch, 1991; Galbicka, Kautz, & Ritch, 1992). This differentiated responding has never achieved a higher overall rate or probability of reinforcement, and by definition has required more responses per reinforcer. These characteristics remain true at all levels of aggregation examined, from different conditions, over sessions, or in blocks of as few as 20 trials. Adding these results to those obtained with other percentile procedures that differentiated response dimensions from interresponse-time duration (e.g., Alleman & Platt, 1973; Arbuckle & Lattal, 1992; Galbicka & Platt, 1986; Kuch & Platt, 1976), to response or changeover duration (e.g., Platt, 1984), to spatial response location (e.g., Galbicka & Platt, 1989; Scott & Platt, 1985), to response variability (Machado, 1989), maintaining either a constant overall reinforcement probability or rate throughout, these results present a challenge to models of behavior change that are predicated on changes in aggregate reinforcement rate or probability. This is not to deny that such factors, if varied, produce systematic changes in behavior. But substantial behavior change often occurs in the absence of changes in these reinforcement dimensions, and sometimes, as is the case here, change occurs even despite unfavorable changes in reinforcement density. The present results indicate that aggregate relations should not be considered fundamental in the control of behavior. Rather, they probably represent the combined effects of more local relations that drive behavior change. Although it was reasonable to begin attempting to quantify behavior by eliminating sources of local variation and developing models of the relatively homogeneous behavior that results (like responding under constant-probability variable-interval schedules), a complete model of behavior must ultimately be able to account for behavior change that is produced both by changes in overall reinforcement rates and in more local relations like the one programmed by percentile schedules. Perhaps it is time to change strategies and attempt to model the local dynamics of responding as they are related to local reinforcement characteristics, while keeping as a linchpin of any such model the requirement that it track the behavioral effects of changing aggregate reinforcement parameters as well.

The present study is meant more to provoke such a local analysis than to provide one. Recent forays into behavioral dynamics, including models based on the sequential structure of responding (e.g., Hoyert, 1992; Palya, 1992) or on linear-systems analysis (e.g., McDowell, Bass, & Kessel, 1992) suggest potential starts. That subjects are capable of discriminating sequential structure in environmental events as well as in behavior should come as no surprise—the areas of psychophysics dealing with topics such as timing (e.g., Gibbon & Allan, 1984), numerosity (e.g., Gallistel, 1989), and so forth are replete with such demonstrations. In fact, the anchoring of behavior around temporal, numerical, spatial, or other cues differentially correlated with reinforcement is so pervasive that models incapable of providing for such correlation must be considered incomplete at best. A viable model of operant behavior must account for the development of behavioral structure as it is warped by reinforcement and the environmental events that act as signposts for biologically significant consequences (cf. Killeen, 1992).

Differentiating response number under targeted percentile schedules may reveal a greater role for sequential dependencies in run length because, unlike traditional reinforcement schedules, percentile schedules are explicitly designed to operate on local structure in responding. Paradoxically, percentile schedules keep the overall probability of reinforcement constant by providing a maximal transition in reinforcement probability (from 0 to 1) for behavior relatively closer to the target. Because the reinforcement contingency is based on the relation between current and recent behavior, it would not be surprising to find a greater degree of sequential structure in behavior than that reported under more typical free-operant arrangements (e.g., Palya, 1992; Peavey, McDowell, & Kessell, 1992). The oscillatory patterns in run length here, for example, suggest some very long-term sequential structure with at least some subjects.

At the other extreme, the data on the deviations presented in Figure 5 suggest a differential result on a trial-by-trial basis depending on the outcome of responding, in that deviations following food presentation were generally more likely to be negative, whereas those following trials without food were more often positive. This contrasts with the data on deviations presented for spatial response location on a circular dimension (Galbicka & Platt, 1989), where deviations were generally centered on the previous response location, with minimal dispersion on trials following food and greater dispersion on those following no food. Both sets of data suggest that reinforcement increases the probability of emitting the behavior most recently associated with food. In the spatial situation, this involves returning to the previous location; here, it involves pressing the right lever, but that in turn means prematurely ending the current run.

This analysis emphasizes that acquisition of differentiated runs requires acquisition and extinction of several, sometimes opposing, operants, and a dynamic model of such acquisition should make this explicit. First, subjects must learn to press the right lever, because responses there terminate the trial and are most closely followed by food. But pressing the right lever alone must ultimately undergo extinction, because only right-lever presses following at least one left-lever press produce any consequences. So left-lever pressing is differentially reinforced and increases. But there are upper limits to the amount of left-lever pressing, imposed both by the percentile schedule, which begins reinforcing shorter runs differentially as comparison runs become increasingly long above the target, and by the inherent delay to food or increased effort involved in completing a longer run. The tendency for runs to stabilize asymmetrically at values slightly below the target most likely reflects the opposition of the differential reinforcement provided by the percentile schedule with that associated with completing a run (cf. Platt, 1984).

There remain higher order dynamics that might differentiate even more complex operants in the present situation. The percentile schedule provides reinforcement differentially for deviations towards a target, not for a particular run per se, and may therefore establish *deviation* as an operant. Hence, a model of behavior in the present study might need to consider not only the run reinforced on a particular trial but also the directional change in behavior from trial to trial when reinforcement was delivered. Similar suggestions have been offered in the past: Skinner (1938) suggested that response number within a fixed interval could be differentiated, Zeiler and colleagues demonstrated that the time to complete a fixed ratio is an operant (see Zeiler, 1977, for a review), and Silberberg and Ziriax (1982) suggested that concurrent-schedule performance is best understood not in terms of individual key pecks but in terms of the differential reinforcement provided for changing between schedules. These suggestions all emphasize that aspects of behavior other than single presses can be conditioned; the percentile procedure used here makes this even more evident by establishing reinforcement contingencies for appropriate deviation.

The present results, therefore, pose a quandary to existing quantitative models of operant behavior. These models presume that behavior matches, maximizes, or is otherwise controlled by some aspect of aggregate reinforcement parameters that yield some overall benefit to the subject (or at least do not worsen its lot). Yet it is difficult to see how behavior of the subjects in the present study could be construed as providing any benefit, except in the short term (i.e., on the next trial). The percentile schedule used here drives aggregate reinforcement parameters away from any long-term optimum, in a sense by placing long-term and short-term goals, or aggregate versus local reinforcement, in opposition. It makes it difficult for subjects to keep doing what they were doing under baseline by offering an immediate incentive for doing something different (i.e., repeating a run that currently dominates the memory will be reinforced with probability w, but moving one step closer to the target will always produce reinforcement). The percentile schedule is, in effect, a socialist version of capitalism realized, in that it guarantees a fixed probability of reinforcement independent of performance while at the same time providing incentive for behavior change. (My thanks to G. Jean Kant for this interesting analogy.) Although overall reinforcement probability remains constant, the promise of reinforcement *on the next trial* drives continuous improvement. Prosperity remains ever around the corner, yet never appears. Viewed in this way, differentiated responding represents a lack of self-control, in that succumbing to local reinforcement contingencies drives overall reinforcement density down (cf. Logue, 1988). But it could as easily be argued that differentiated responding *demonstrates* self-control, and not a lack thereof, because right-lever presses must increasingly be delayed for reinforcement on the next trial. Therein lies the quandary.

REFERENCES

Alleman, H. D., & Platt, J. R. (1973). Differential reinforcement of interresponse times with controlled probability of reinforcement per response. *Learning and Motivation*, **4,** 40–73.

Arbuckle, J. L., & Lattal, K. A. (1992). Molecular contingencies in schedules of intermittent punishment. *Journal of the Experimental Analysis of Behavior*, **58,** 361–375.

Davison, M., & McCarthy, D. (1988). *The matching law: A research review.* Hillsdale, NJ: Erlbaum.

Galbicka, G. (1988). Differentiating *The Behavior of Organisms. Journal of the Experimental Analysis of Behavior*, **50,** 343–354.

Galbicka, G., Fowler, K. P., & Ritch, Z. J. (1991). Control over response number by a targeted percentile schedule: Reinforcement loss and the acute effects of *d*-amphetamine. *Journal of the Experimental Analysis of Behavior*, **56,** 205–215.

Galbicka, G., Kautz, M. A., & Ritch, Z. J. (1992). Reinforcement loss and behavioral tolerance to *d*-amphetamine: Using percentile schedules to control reinforcement density. *Behavioural Pharmacology*, **3,** 535–544.

Galbicka, G., & Platt, J. R. (1986). Parametric manipulation of interresponse-time contingency independent of reinforcement rate. *Journal of Experimental Psychology: Animal Behavior Processes*, **12,** 371–380.

Galbicka, G., & Platt, J. R. (1989). Response-reinforcer contingency and spatially defined operants: Testing an invariance property of phi. *Journal of the Experimental Analysis of Behavior*, **51,** 145–162.

Gallistel, C. R. (1989). Animal cognition: The representation of space, time and number. *Annual Review of Psychology*, **40,** 155–189.

Gibbon, J., & Allan, L. (Eds.). (1984). *Timing and time perception.* New York: New York Academy of Sciences.

Hoyert, M. S. (1992). Order and chaos in fixed-interval schedules of reinforcement. *Journal of the Experimental Analysis of Behavior*, **57,** 339–363.

Hursh, S. R. (1980). Economic concepts for the analysis of behavior. *Journal of the Experimental Analysis of Behavior*, **34,** 219–238.

Killeen, P. R. (1992). Mechanics of the animate. *Journal of the Experimental Analysis of Behavior*, **57,** 429–463.

Kuch, D. O., & Platt, J. R. (1976). Reinforcement rate and interresponse time differentiation. *Journal of the Experimental Analysis of Behavior*, **26,** 471–486.

Logue, A. W. (1988). Research on self-control: An integrating framework. *Behavioral and Brain Sciences*, **11,** 665–709. (includes commentary)

Machado, A. (1989). Operant conditioning of behavioral variability using a percentile reinforcement schedule. *Journal of the Experimental Analysis of Behavior*, **52,** 155–166.

McDowell, J. J, Bass, R., & Kessel, R. (1992). Applying linear systems analysis to dynamic behavior. *Journal of the Experimental Analysis of Behavior*, **57,** 377–391.

Palya, W. L. (1992). Dynamics in the fine structure of schedule-controlled behavior. *Journal of the Experimental Analysis of Behavior*, **57,** 267–287.

Peavey, M. E., McDowell, J. J. & Kessel, R. (1992). Shaw's stored information as a quantitative measure of sequential structure. *Behavior Research Methods, Instruments, & Computers*, **24,** 228–237.

Platt, J. R. (1973). Percentile reinforcement: Paradigms for experimental analysis of response shaping. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in theory and research* (Vol. 7, pp. 271–296). New York: Academic Press.

Platt, J. R. (1984). Motivational and response factors in temporal differentiation. In J. Gibbon & L. Allan (Eds.), *Timing and time perception* (pp. 200–210). New York: New York Academy of Sciences.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). New York: Appleton-Century-Crofts.

Scott, G. K., & Platt, J. R. (1985). Model of response-reinforcer contingency. *Journal of Experimental Psychology: Animal Behavior Processes*, **11,** 152–171.

Silberberg, A., & Ziriax, J. M. (1982). The interchangeover time as a molecular dependent variable in concurrent schedules. In M. L. Commons, R. J. Herrnstein, & H. Rachlin (Eds.), *Quantitative analyses of behavior: Vol. 2. Matching and maximizing acccounts* (pp. 131–151). Cambridge, MA: Ballinger.

Skinner, B. F. (1938). *The behavior of organisms.* New York: Appleton-Century.

Skinner, B. F. (1966). Operant behavior. In W. K. Honig (Ed.), *Operant behavior: Areas of research and application* (pp. 12–32). New York: Appleton-Century-Crofts.

Snapper, A. G., & Inglis, G. B. (1985). *SKED-11 software system.* Kalamazoo, MI: State Systems, Inc.

Zeiler, M. (1977). Schedules of reinforcement: The controlling variables. In W. K. Honig & J. E. R. Staddon (Eds.), *Handbook of operant behavior* (pp. 201–232). Englewood Cliffs, NJ: Prentice-Hall.

Received September 15, 1992
Final acceptance January 8, 1993