Forming chunks in instrumental learning: The role of overshadowing

E. J. CAPALDI, KIMBERLY M. BIRMINGHAM, and RONALD M. MILLER
Purdue University, West Lafayette, Indiana

Lower order units combined into higher order functional units (e.g., letters into words) are called *chunks*. The hypothesis tested here suggests that chunks are formed when memories of goal events signal other goal events, a signal capacity that can be reduced when memories are seriously overshadowed by other more valid cues called *grouping cues*. In support of this hypothesis, it was found, in each of three experiments employing rats in runways, that the capacity of the memory of nonreward to signal reward depended on its validity relative to that of two other situational cues, runway brightness cues and a change in brightness cues, from one trial to the next. Specifically, the capacity of the memory of nonreward to signal reward was much more seriously reduced when it was in competition with more valid situational cues (.33 vs. 1.00, Experiment 1; .50 vs. 1.00, Experiment 2) than when it was in competition with situational cues having either the same validity (.33, Experiment 1) or a slightly higher validity (.33 vs. .50, Experiment 3). It was also found that grouping each of two series was less effective in producing chunking than was grouping only one of the series.

A critically important task in any area of scientific activity is the identification and empirical isolation of theoretically useful units of analysis that can be used to explain a wide variety of phenomena (e.g., atoms). According to a recent view (Capaldi, 1992), a variety of empirical findings suggest that a common characterization of the instrumental response in many learning theories provides a unit of analysis that is far from optimal, one which, for the most part, should be replaced by another broader or more inclusive unit. Consider first that, in many influential theories ranging historically from Thorndike's (1911) to Hull's (1943) to the present (e.g., Amsel, 1992), the instrumental response is characterized as an intratrial phenomenon, such as a trip from the startbox to the goalbox in a runway. As a concrete example, in Hull's (1943) theory, the strength of the tendency to make such intratrial responses was said to increase when they were reinforced and to decrease when they were nonreinforced. The intratrial conception of the response, despite its general popularity, is difficult to reconcile with a wide variety of empirical findings (see, e.g., Capaldi, 1992, 1994; Capaldi, Miller, Alptekin, & Barry, 1990).

Chunking occurs when lower order elements (e.g., letters) are combined into a new higher order functional unit (e.g., words). The intratrial response as a fundamental unit of analysis was called the *trial chunk* by Capaldi (1992). In a trial chunk, individual elements of a trial (e.g., leaving the startbox, entering the goalbox, approaching the goalcup, etc.) are combined into a single functional unit. The major inadequacy of the trial chunk as a useful

level of theoretical analysis, as revealed by a variety of experimental findings (see, e.g., Capaldi, 1992, 1994; Capaldi et al., 1990), is that, in instrumental situations, animals routinely form trial chunks into a higher order more inclusive unit called the *series chunk*. While rats are capable of forming series chunks into still higher level chunks called *list chunks* (see, e.g., Capaldi, 1992), series chunks seem to be the most commonly employed unit in instrumental situations, at least as such investigations are conducted at present.

In a series chunk, reinforcing events associated with two or more trial chunks are combined into a functional unit. As a specific example, the memory of nonreinforcement from one trial, by being retrieved on a subsequent reinforced trial, may become a signal for reinforcement. In a series chunk of this sort, the animal learns that reinforcement is contingent on the completion of two trips down the runway, a nonreinforced trip followed by a reinforced trip. Thus, series chunks explain, among other things, why, with repeated administration of a series of, say, three successive nonreinforced runway acquisition trials that terminate in a reinforced trial (NNNR), speed of running will come eventually to progressively increase, not decrease, over the successive nonreinforced trials (Capaldi & Verry, 1981). This finding and similar findings appear to be completely paradoxical, employing trial chunks as a unit of analysis (see Capaldi, 1992).

One of the major variables employed to affect chunk formation is how events are grouped. Events presented under some common stimulus condition may come to be grouped together. For example, each of a series of instrumental trials may occur in a black runway or at the same intertrial interval. Single grouping of this sort is assumed to facilitate formation of a single chunk. If a series of events is presented in two groups by, for example,

Correspondence should be addressed to E. J. Capaldi, Department of Psychological Sciences, 1364 Psychology Bldg., Purdue University, West Lafayette, IN 47907-1364 (e-mail: julie@psych.purdue.edu).

interposing some distinctive cue in the middle of the series, two separate chunks may be formed, one before and one after the distinctive interpolated cue. Currently, no generally accepted theory of how grouping affects chunk formation is available.

The purpose of the three experiments reported here was to test a theory, suggested by Capaldi, Verry, Nawrocki, and Miller (1984), of how grouping cues affect chunk formation in situations in which reinforcement events (reward, nonreward, etc.) are presented successively in a fixed and, thus, predictable order. This theory has yet to be tested directly: it was simply suggested as a plausible explanation of available grouping investigations by Capaldi et al. (1984). According to Capaldi et al., reward events contingent on an instrumental response may be signaled by two different sorts of cues that are in competition with each other. These stimuli are the memory of the reward event stored on the prior trial (item memory) and various situational cues, such as general apparatus cues, visual cues arising from a manipulandum, and so on (see, e.g., Capaldi, 1994; Haggblom, 1981, 1983; Neely & Wagner, 1974). A series chunk consists in one goal event signaling the next goal event. A chunk boundary occurs when a goal event is signaled, not by the memory of the prior reward event but by some other cue. In explicit grouping, the grouping cue may be a more valid signal for reward than either the item memory or various situational cues (called *other cues* here). In this case, the item memory will be more seriously overshadowed by the grouping cue than by the other cues, and the capacity of the item memory to signal reward will be more greatly reduced by the grouping cues than by the other cues.

The overshadowing interpretation of grouping, as will be shown in the General Discussion section, is consistent with the results of animal serial learning experiments employing reward events as items. In those experiments, variations in relative validity occurred, but inadvertently rather than deliberately. Essentially variation in the validity of cues occurred as a consequence of manipulating some other variable. This study appears to be the first in which the relative validities of these cues were deliberately and explicitly varied, and, thus, the relative validities of all cues could be determined. In this study, we explicitly examined the ability of an item memory of a given validity to signal a 10-pellet reward, depending on the validity of other cues and/or explicit grouping cues. Other cues were runway brightness cues. The grouping cue was a change in runway brightness, always from Trial 2 to Trial 3.

How and when grouping cues may overshadow item memories in an instrumental situation may be illustrated by a specific example. In a runway of a given brightness (black on half the days, white on the other half) in which successive trials always terminate in 0-, 0-, and 10-pellet reward (single-group condition), the memory of nonreward has a validity of .50, since it is retrieved on Trial 2 (a nonrewarded trial) and on Trial 3 (a rewarded trial). The validity of each of the brightness cues, which com-

pete with the memory of nonreward as a signal for 10 pellets, is .33, since only one third of the trials is reinforced. Of course, a variety of other cues (e.g., being picked up by the experimenter, etc.) have a validity of .33. In order to simplify exposition, such cues are ignored in this report. In any event, introducing a grouping cue, a change in runway brightness from Trial 2 to Trial 3 (two-group condition), would have the following effects: First, it would not alter the validity of the 0-pellet memory because, as data reported by Capaldi et al. (1984) indicate, the introduction of a grouping cue in extinction elevated running speed, which would not have been possible in the absence of memory retrieval. In any event, the brightness-change grouping cue, by being distinctive and consistently signaling 10 pellets, has a validity of 1.00, and so it can more effectively overshadow the memory of 0 pellets as a signal for 10 pellets than can a brightness cue having a validity of .33. Thus, the 0-pellet item memory will have a stronger capacity to signal 10 pellets in the single-group condition than in the two-group condition. Accordingly, it may be expected that running speed on Trial 2 of the 0-0-10 series, which is signaled by the 0-pellet memory, will be more rapid in the single-group condition than in the two-group condition.

As is perhaps clear from the example provided, the validity of item memories is reduced when either the same item memories or similar item memories signal different items. In each experiment each day, rats in a runway received fixed series of reinforced and nonreinforced trials. Of primary concern in each experiment was the ability of the animals to correctly predict nonreinforcement on certain critical trials as indicated by slow running. In single-group conditions, each day, all trials occurred in a runway of the same brightness (black on half the days, white on the other half). In two-group conditions, a change in runway brightness was the grouping cue: Runway brightness changed from Trial 2 to Trial 3, from black to white on half the days, and from white to black on the other half. In each of the three experiments reported here, the only respect in which the two groups differed was whether runway brightness did or did not change from Trial 2 to Trial 3. In Experiment 1, the validity of the grouping cue in the two-group condition, which was 1.00, was much higher than that of either the item memory (.33) or the other cues (.33). Thus, in the two-group condition, it was expected that the item memory would be seriously overshadowed by the grouping cue. In the single-group condition of Experiment 1, the capacity of the item memory to signal reward was expected to be greater than in the two-group condition because its validity equaled that of the other cues (both .33).

In Experiment 2, both of the groups that were employed received a grouping cue, but in only one of the groups was it more valid than the item memory (1.00 vs. .50); in the other group, the item memory and the grouping cue were equally valid (both .50). In Experiment 3, as in Experiment 1, the validity of the item memory and the other cues was .33. However, in the two-group condition, the validity of the grouping cue was .50 rather than 1.00.

Determining the extent to which an item memory having a validity of .33 might be more seriously overshadowed by a grouping cue having a validity of .50 than by other cues having a validity of .33 became a relevant question given the findings obtained in Experiments 1 and 2. An aspect of Experiments 2 and 3 that deserves mention (but will not receive further comment until the General Discussion section) is that one of the groups in each of those experiments had both its series grouped rather than just one series. On the basis of available findings, grouping both series, as in Experiments 2 and 3, might be thought to facilitate chunking; however, according to the overshadowing view, it would actually reduce chunking.

EXPERIMENT 1

On each day of acquisition in Experiment 1, the two groups compared received a series of three runway trials (A), a series of two runway trials (B), and a single runway trial (C). The interval elapsing between the trials of a series was about 15 sec. The interval between the A, B, and C events, presented in different orders over days, was about 10 min. In the three-trial series, the first two trials terminated in nonreward, with Trial 3 terminating in 10-pellet reward (referred to as a 0–0–10 series). Both trials of the two-trial series were nonrewarded (referred to as a 0–0 series). The single trial (C) always terminated in 10 pellets. In designating the groups, the first number in the group name denotes the validity of the brightness cue or the brightness-change cue. The second number in the group name denotes the validity of the item memory. In Group .33/.33A (a .33/.33 group, Group .33/.33B was employed in Experiment 3), the single-group condition, all trials occurred each day in a runway of a single brightness (black on half the days, white on the other half). Group 1.00/.33, the two-group condition, was treated exactly as was Group .33/.33A, with one exception: Runway brightness changed from Trial 2 to Trial 3 in the 0–0–10 series (either from white to black or from black to white). Over a 2-day block, the validity of each of the competing brightness cues was .33 in both groups, as was the validity of the item memory of 0 pellets (.50 in the 0–0–10 series, and .00 in the 0–0 series). However, in Group 1.00/.33, unlike in Group .33/.33A, a highly valid cue (1.00)—the change in alley brightness—signaled 10 pellets. Thus, the capacity of the item memory of 0 pellets to signal 10 pellets should be more seriously overshadowed in Group 1.00/.33 than in Group .33/.33A; therefore, on Trial 2 of both the 0–0–10 series and the 0–0 series (signaled by the memory of 0 pellets), Group 1.00/.33 should run more slowly than Group .33/.33A.

Method

Subjects. The subjects were 8 naive male albino rats (virus-free) about 70 days old when purchased from the Holtzman Co., Madison, WI, and about 80 days old at the start of experimental training.

Apparatus. The apparatus consisted of side-by-side runways, one black, the other white. The runways were 194.94 cm long, 8.8 cm wide, enclosed by 12.7-cm sides covered by a wire mesh on a hinged

frame. The startbox of each runway was 23.3 cm long. The goalbox of each runway was 36.15 cm long. Raising the startbox door started a timer in a computer (IBM PC XT), which was stopped when the rat broke a photobeam located 158.13 cm beyond the startbox door. Pellets (45 mg, Noyes) were placed in a goalcup, which was 4.0 cm in diameter and 1.5 cm in depth, cut into a $6.5 \times 8.8 \times 4.5$ cm block of wood. A brass door was lowered, confining the rats to the goalbox.

Pretraining. The rats were caged individually and given ad-lib food and water for 7 days after arriving at the laboratory. On Day 8, pretraining began. On Day 1 of pretraining, food deprivation began consisting of 18 g of Laboratory Rodent Diet 5001 (PMI Nutrient International) each day minus any amount eaten in the runway. On Days 1–7 of pretraining, the rats were handled. On Days 8–11 of pretraining, runway exploration occurred. On Day 8, the rats were fed ten 0.045-g Noyes pellets in the home cage; on Days 8–11, they were allowed 3 min to explore the runways, which had 10 pellets scattered about. The startbox and goalbox doors were lowered and raised at about the 1.5-min mark in order to habituate the rats to the noise. On Days 8, 9, and 10 of pretraining, the rats explored the white, black, and white runways, respectively.

Experimental training. After 10 days of pretraining, experimental training began. There were two groups of 4 rats each. All rats were brought into the experimental room from the colony room in their home cages, where they spent the interval between trials of a series and the interval between series. Each day, each rat received two different series of trials and a single trial. A trial began when the rat was placed into the startbox. The startbox door was opened. If the rat did not enter the goalbox within 30 sec, it was picked up and placed there to receive the scheduled event. On nonrewarded trials, the goalbox was false-baited by going through the exact motions as on a reinforced trial. One series was 0–0–10, the other series was 0–0, and the single trial was 10 (note that 0 refers to nonreward consisting of 15-sec confinement in the empty goalbox; 10 refers to ten 0.045-g Noyes pellets). Trials of a series were separated by an interval of about 15 sec; series were separated by an interval of about 10–15 min. Referring to the 0–0–10 series as A, the 0–0 series as B, and the single 10-pellet trial as C, six different orders of the series were used, each occurring once in each 6-day block: ABC, CAB, BCA, CBA, ACB, and BAC.

The order in which the rats were run was randomly determined each day. Each rat received all trials of its three-item or two-item series before the next rat received its series. Group .33/.33A received all of its trials in the black runway on odd days and in the white runway on even days. Group 1.00/.33 was treated similarly, but with one exception: Runway brightness changed from black to white on odd days and from white to black on even days from Trial 2 to Trial 3 of the 0–0–10 series. After the last rat was run, all 8 rats were returned to the colony room, where they were housed in individual cages and fed the daily ration about 10 min later.

Results and Discussion

In this report, differences are said to be significant only if $p < .05$. Figure 1 shows mean median running speeds on each of the trials of the 0–0–10 and 0–0 series in blocks of 4 days for Group 1.00/.33 and Group .33/.33A. Running speeds on the single 10-pellet trial are not shown in Figure 1. Running speeds on that trial were nearly identical for the two groups. On the last block of trials, for example, the speeds were 104.49 and 105.74 cm/sec, for Groups 1.00/.33 and .33/.33A, respectively.

Figure 1 shows that, over the initial blocks of trials, Groups 1.00/.33 and .33/.33A differed little. By the last two blocks of trials, Group 1.00/.33, but not Group .33/.33A, showed a marked tendency to correctly anticipate nonre-

Figure 1. Mean median running speed on each trial of the 0-0-10 and 0-0 series in blocks of 4 days for Group 1.00/.33 and Group .33/.33A.

ward by running slowly on Trial 2 of both the 0-0-10 and the 0-0 series. An analysis of variance (ANOVA) applied to the data shown in Figure 1, including the single 10-pellet trial, indicated that significant differences were associated with both the group $\times$ trials [$F(5,10) = 3.71$] and the group $\times$ trials $\times$ blocks [$F(70,140) = 5.07$] interactions. These interactions, it is clear, were due entirely to slower running by Group 1.00/.33 than by Group .33/.33A on Trial 2 of the 0-0-10 series and on Trial 2 of the 0-0 series. For Group 1.00/.33, the standard error of the means on each of the Blocks 1-7 were, respectively, 7.21, 9.92, 6.98, 5.52, 6.20, 9.55, and 8.26. For Group .33/.33A, they were 9.56, 11.77, 9.72, 8.87, 9.78, 10.27, and 10.17.

In Experiment 1, each day, both groups received a 0-0-10 series, a 0-0 series, and a single 10-pellet trial. In Group 1.00/.33, but not in Group .33/.33A, a distinctive grouping cue—a change in runway brightness—reliably signaled the 10-pellet trial of the 0-0-10 series. It was found that, by the end of training, Group 1.00/.33 ran much more slowly than Group .33/.33A on Trial 2 of both the 0-0-10 and the 0-0 series. Indeed, by the end of training, some 28 days, Group .33/.33A did not even begin to run slowly on Trial 2 of the 0-0-10 and the 0-0 series. The present explanation of this finding is as follows: On Trial 3 of the 0-0-10 series, two competing cues signaled 10 pellets. One was the memory of the prior 0-pellet item, the competing cues being a brightness cue in Group .33/.33A and a change in runway brightness in Group 1.00/.33. In Group .33/.33A, the validity of the competing cue was the same as that of the 0-pellet item memory (.33). But, in Group 1.00/.33, the validity of the competing cue was much higher than that of the 0-pellet item memory (1.00 vs. .33). Consequently, the item memory as a signal for 10 pellets was much more effectively overshadowed in Group 1.00/.33 than in Group .33/.33A. Thus, when the 0-pellet memory was retrieved on Trial 2 of the 0-0-10 series and on Trial 2

of the 0-0 series, it elicited much stronger responding in Group .33/.33A than in Group 1.00/.33.

EXPERIMENT 2

Experiment 2 was in many respects similar to Experiment 1, with notable exceptions mentioned below. In Experiment 2, each day, all rats received a single presentation of two somewhat different series: a 10-0-10 series and a 10-0-0 series. Each day, for half the rats, the 10-0-10 series occurred first, with the 10-0-0 series occurring about 10 min later. The remaining rats received the two series in the opposite order. Previous investigations have shown that when the 10-0-10 and 10-0-0 series are presented in the same order each day, rats are able to correctly anticipate all reward outcomes, running faster on 10-pellet trials than on 0-pellet trials (see, e.g., Capaldi, Nawrocki, & Verry, 1983).

In Experiment 2, we used two groups, both of which received as a grouping cue a change in runway brightness from Trial 2 to Trial 3, black to white on half the days and white to black on the other half. In Group 1.00/.50, the grouping cue occurred in only one of the series, the 10-0-10 series; in Group .50/.50, it occurred in both series. In both groups, the validity of the 0-pellet item memory was .50, as was the validity of the brightness cue. However, the validity of the grouping cue, the change in runway brightness, was 1.00 in Group 1.00/.50, since it occurred only prior to 10-pellet trials; it was .50 in Group .50/.50, since it occurred prior to both 10-pellet and 0-pellet trials.

Method

Subjects. The subjects were 8 rats of the same description as in Experiment 1.

Apparatus. The apparatus was the same as that employed in Experiment 1.

Pretraining. Pretraining was identical to that employed in Experiment 1.

Experimental training. Experimental training in Experiment 2 was identical to that employed in Experiment 1, except for the following differences. Each rat received two series each day: a 10-0-10 series and a 10-0-0 series. Each rat received the series in a given order each day. Referring to the 10-0-10 series as A and the 10-0-0 series as B, 2 rats in each group received the A series first and the B series second each day; the other 2 rats in the group received the B series followed by the A series. In Group 1.00/.50, runway brightness changed from black to white or from white to black from Trial 2 to Trial 3 of the 10-0-10 series. In Group .50/.50, the change in runway brightness occurred from Trial 2 to Trial 3 in both its 10-0-10 series and its 10-0-0 series.

Results and Discussion

Figure 2 shows mean median running speeds on each of the trials of the 10-0-10 and 10-0-0 series in blocks of 4 days for Group 1.00/.50 and Group .50/.50. As is clearly evident on the last block of trials, Group 1.00/.50 showed, throughout training, a much greater tendency than Group .50/.50 to correctly anticipate 10-pellet and

Figure 2. Mean median running speed on each of the trials of the 10-0-10 and 10-0-0 series in blocks of 4 days for Group 1.00/.50 and Group .50/.50.

0-pellet trials by running faster to the former than to the latter. An ANOVA applied to the data shown in Figure 2 revealed that significant differences were associated with group [$F(1,6) = 37.08$] and with the group $\times$ trials interaction [$F(5,30) = 4.27$]. A breakdown of the interaction employing Newman-Keuls tests indicated that Group 1.00/.50 was better able to correctly anticipate 10-pellet and 0-pellet trials than was Group .50/.50. First, three critical between-groups differences were significant. Group 1.00/.50 ran more slowly than Group .50/.50 on Trial 2 of the 10-0-10 series, on Trial 2 of the 10-0-0 series and on Trial 3 of the 10-0-0 series. Only one within-groups difference was significant: Group 1.00/.50 ran more slowly on Trial 3 of the 10-0-0 series than on Trial 3 of the 10-0-10 series. For Group 1.00/.50, the standard error of the means on each of the Blocks 1-7 were, respectively, 6.09, 10.83, 8.30, 9.50, 9.03, 8.33, and 8.61. For Group .50/.50, they were 3.60, 6.52, 7.42, 7.77, 7.59, 8.02, and 8.26.

In Experiment 2, Group 1.00/.50 ran more slowly on Trial 3 of the 10-0-0 series than did Group .50/.50. The overshadowing explanation of this finding is straightforward. In the 10-0-10 series, the 10-pellet item on Trial 3 was signaled by two cues: the memory of nonreward and the change in runway brightness. The validity of the memory of nonreward was .50 in both groups. But the grouping cue, the change in runway brightness, was more valid in Group 1.00/.50 (1.00) than in Group .50/.50 (.50). This resulted in the memory of nonreward as a cue for 10 pellets being more seriously overshadowed by the competing grouping cue in Group 1.00/.50 than in Group .50/.50. Thus, the memory of nonreward, which was retrieved on Trial 3 of the 10-0-0 series, elicited weaker responding on that trial in Group 1.00/.50 than in Group .50/.50.

It was also found in Experiment 2 that Group 1.00/.50 ran more slowly on Trial 2 of both the 10-0-10 and the 10-0-0 series than did Group .50/.50. The memory signaling 0 pellets on Trial 2 was that of reward. It appears that the memory of reward had a weaker capacity to signal reward in Group 1.00/.50 than in Group .50/.50. This finding, which is not of direct concern here, is nevertheless interesting. A possible explanation of the finding is that elements of the memory of reward were retrieved on Trial 3 of both series and so were more effectively overshadowed as signals for reward in Group 1.00/.50 than in Group .50/.50.

EXPERIMENT 3

In Experiment 1, an item memory having a validity of .33 was more effectively overshadowed by a grouping cue having a validity of 1.00 than by other cues having a validity of .33. In Experiment 2, an item memory having a validity of .50 was more effectively overshadowed by a grouping cue having a validity of 1.00 than by a grouping cue having a validity of .50. In sum, in Experiments 1 and 2, grouping cues were shown to be quite effective in overshadowing item memories when their validity was very much higher than that of the item memories and other cues. However, the results of Experiments 1 and 2 do not rule out the possibility that grouping per se may be efficacious in producing chunks, even when the validity of the grouping cue is equal to or slightly higher than that of the item memory and other cues. A finding of that sort would be difficult to understand from the viewpoint of the present overshadowing view.

In Experiment 3, each day, each of two groups received a 10-0-10 series and a 0-0-0 series in different orders over days. In the single-group condition (Group .33/.33B),

runway brightness changed from black to white on half the days and from white to black on the other half, and, so, as in Experiment 1, the validity of the brightness cues was .33, as was the validity of the 0-pellet memory. In the two-group condition (Group .50/.33), runway brightness changed from Trial 2 to Trial 3 in both the 10-0-10 series and the 0-0-0 series. Thus, in Group .50/.33, the validity of the grouping cue, which was .50, was only slightly higher than that of the 0-pellet item memory and the other cues, both of which were .33.

Method

Subjects. The subjects were 8 rats of the same description as employed in Experiments 1 and 2.

Apparatus. The apparatus was the same as that employed in Experiments 1 and 2.

Pretraining. Pretraining was identical to that employed in Experiments 1 and 2.

Experimental training. Experimental training was identical to that employed in Experiments 1 and 2, with the following exceptions. Each day, each rat received a 10-0-10 series and a 0-0-0 series. Within each 6-day block, the series were given in the order AB, BA, BA, AB, BA, and AB, respectively. Group .33/.33B received all of its trials in a black runway on odd days and in a white runway on even days. Group .50/.33 was trained similarly, except that runway brightness changed from black to white on odd days and from white to black on even days from Trial 2 to Trial 3 of both the 10-0-10 series and the 0-0-0 series.

Results and Discussion

Figure 3 shows mean median running speeds on each of the trials of the 10-0-10 and 0-0-0 series in blocks of 4 days for Group .33/.33B and Group .50/.33. In the early blocks, the two groups ran at about the same speed on all trials. By the last two blocks of trials, Group .33/.33B on an absolute basis ran more slowly than did Group .50/.33. In any event, by the last block of trials, both groups were

running more slowly on 0-pellet trials than on 10-pellet trials. An ANOVA applied to the data shown in Figure 3 revealed the following. No differences involving groups was significant. Significant differences were associated with blocks [$F(8,16) = 34.03$] and with trials [$F(5,100) = 4.42$]. The significant trials effect simply indicates that running speeds were greater on 10-pellet trials than on 0-pellet trials. Since it appeared that, in Block 9, the animals may have been running slower on Trial 3 of the 0-0-0 series, a separate ANOVA was applied to the data of Block 9. No significant effects were found. For Group .50/.33, the standard error of the means on each of the Blocks 1-9 were, respectively, 9.04, 11.02, 10.76, 9.98, 10.35, 11.25, 11.76, 12.69, and 10.74. For Group .33/.33B, they were 6.62, 9.05, 9.82, 11.57, 10.13, 10.99, 11.18, 10.53, and 10.56.

Experiments 1 and 2 demonstrated that an item memory having a given validity would be much more seriously overshadowed by a cue having a much higher validity than by a cue having the same validity. Experiment 3 demonstrated that an item memory having a validity of .33 was no more greatly overshadowed by a grouping cue having a slightly higher validity (.50) than by other cues having the same validity (.33). Thus, Experiment 3 demonstrated that it is not grouping per se that facilitated chunk formation but rather the extent to which the validity of the grouping cue was higher than that of other cues and item cues. Although it is possible that, if training had continued, Trial 3 differences may have emerged. Militating against this conclusion, at least to some extent, is the fact that differences between the groups developed in Experiments 1 and 2 following many fewer training trials than were employed in Experiment 3.

Given the findings obtained in Experiments 1 and 2, it seems clear that the anticipation of 0 pellets in Group .50/.33 would have been better if only the 10-0-10 series had been grouped rather than both the 10-0-10 and the 0-0-0 series.

GENERAL DISCUSSION

A seemingly important finding obtained in Experiments 2 and 3, mentioned in passing earlier, deserves discussion. Previous investigations have shown that accuracy in anticipating an item may be improved by dividing longer series into shorter ones via grouping (see, e.g., Wickelgren, 1964). This was certainly the case in Experiment 1, in which 0-pellet items were much better anticipated by Group 1.00/.33 than by Group .33/.33A. However, as shown in Experiments 2 and 3, grouping longer series into shorter ones is efficacious only to the extent that the validity of the grouping cue is higher relative to that of item memories and other cues. Thus, in Experiment 2, anticipation of 0 pellets was better when only one series (Group 1.00/.50) rather than both series (Group .50/.50) were divided into shorter series. Moreover, as indicated, grouping both series rather than only the 10-0-10 series in Experiment 3 may be assumed to

Figure 3. Mean median running speed on each of the trials of the 10-0-10 and 0-0-0 series in blocks of 4 days for Group .33/.33B and Group .33/.50.

have reduced correct anticipation of 0-pellet trials in Group .50/.33. Poorer anticipation of items, when both series rather than only one is grouped as in Experiments 2 and 3, may seem paradoxical from some standpoints given earlier reported findings, such as those of Wickelgren (1964). However, that result is entirely understandable in terms of the present overshadowing view, since grouping both series reduced the validity of the grouping cue.

The finding described above and other findings obtained here highlight an important implication of the present overshadowing view—namely, that grouped series and ungrouped series are similar in an important respect. That is, item memories may be in competition with both “sorts” of stimuli. At bottom, the question is, How valid are these competing stimuli relative to the item memories? Thus, in dealing with any series, the important matter is, How valid are the item memories relative to competing cues whether these are other cues or grouping cues? The present results seem to supply strong support for this proposition.

In Experiment 1, a 0-pellet item memory having a validity of .33 was much more seriously overshadowed by a grouping cue having a validity of 1.00 than by other cues having a validity of .33. In Experiment 2, an item memory having a validity of .50 was much more seriously overshadowed by a grouping cue having a validity of 1.00 than by a grouping cue having a validity of .50. In Experiment 3, an item memory having a validity of .33 did not appear to be more seriously overshadowed by a grouping cue having a validity of .50 than by other cues having a validity of .33. Taken together, these findings suggest that the important matter is not whether or how a series is grouped but how valid are competing cues (grouping cues or other cues) relative to the item memories.

As indicated, in previously reported runway grouping investigations, the relative validity of cues was varied inadvertently rather than as a result of deliberate intention. Common characteristics of the most important of the findings obtained in those investigations are as follows: A 0-pellet item is preceded by a small reward item (e.g., 2 pellets, 1 pellet, or 0 pellets) and is followed by a larger reward item (e.g., 10 pellets or 14 pellets). Of primary concern in such investigations was the development of slow running on the 0-pellet trial. This has been shown to depend on how repetitions of the series are grouped. For example, if a distinctive grouping cue signals the 14-pellet item in repetitions of the five item series 14–7–3–1–0, running is slow to the 0-pellet item. But if the successive series are ungrouped or if the distinctive grouping cue occurs elsewhere in the series, say, between the 1-pellet item and 0-pellet item, running is rapid on the 0-pellet trials. Such findings have been obtained not only for the 14–7–3–1–0 series but for a variety of other series (e.g., 10–2–0–10; see, e.g., Capaldi, Nawrocki, Miller, & Verry, 1986; Capaldi et al., 1984; Fountain, Henne, & Hulse, 1984).

The above-described findings are consistent with the present view. Where a distinctive and, thus, highly pre-

dictive grouping cue precedes the large reward item, the 0-pellet memory will be seriously overshadowed as a signal for larger reward, and so the small reward item that precedes and signals 0 pellets will elicit slow running. Where the distinctive grouping cue does not signal large reward, either because repetitions of the series are not grouped or because the grouping cue occurs elsewhere in the series, the large reward item will occur in connection with cues of relatively low validity. Thus, the 0-pellet memory will be less overshadowed on the large reward trial and so will acquire a relatively strong capacity to signal large reward. This relatively strong signal capacity will generalize to the highly similar small reward item, which precedes the 0-pellet item. Thus, the memory of the small reward item will elicit strong responding on the 0-pellet trial. The general analysis supplied above is consistent with a wide variety of runway grouping findings that have been reported. Also, with slight modification, the analysis will explain findings obtained using other sorts of series (e.g., repetition of 0–1–15 series; Capaldi et al., 1986).

The serial learning situation employed here is a conventional type first examined experimentally by Ebbinghaus (1885). In the present type of serial learning tasks, items are presented successively in a fixed order, the subject's task being to correctly anticipate each item prior to its occurrence. Two other noteworthy situations have been employed to study chunk formation in animals. In one situation, extensively employed by Terrace and colleagues (see, e.g., Terrace, *in press*), pigeons or monkeys are required to respond in a particular order to items that are displayed simultaneously. A variety of differences between series presented simultaneously versus successively have been detailed by Terrace (*in press*). In another situation, rats in a radial maze received different types of food items, each type occurring in the same array over successive trials (e.g., Macuda & Roberts, 1995). In each of the situations described above a variety of evidence for chunking has been presented. In the Terrace situation, for example, pigeons learned ungrouped series (e.g., all five items were colors) more slowly than grouped series (e.g., three items were colors, two items were forms). In the radial maze situation, for example, rats might enter all arms containing Food Item A first, then all arms containing Food Item B next, and, finally, all arms containing Food Item C.

The extent to which the chunks isolated in the three situations described above are similar is at this time a matter of conjecture. Important as that matter may be, it is not of primary concern here. Our ultimate objective, mentioned earlier, is to isolate and identify chunks that are useful for explaining a variety of conventional instrumental learning phenomena. The series chunk seems especially useful in this regard (see, e.g., Capaldi, 1992, 1994). The experimental findings reported here appear to describe conditions that govern the formation of series chunks in conventional instrumental learning situations. As we come to better understand those conditions, we

will be better able to determine in particular instrumental tasks what sorts of chunks are formed and, thus, what level of theoretical analysis is required: one in terms of lower level trial chunks, one in terms of higher level series chunks, or perhaps even one in terms of higher level list chunks mentioned earlier. However, as instrumental learning investigations are conducted at present, series chunk formation is common. Typically, in instrumental investigations, a series of two or more intratrial responses (bar-presses, trips down a runway, etc.) is allowed to occur in connection with some common stimuli (e.g., all trials at the same intertrial interval), a circumstance that favors series chunk formation.

REFERENCES

- AMSEL, A. (1992). *Frustration theory: An analysis of dispositional learning and memory*. New York: Cambridge University Press.
- CAPALDI, E. J. (1992). Levels of organized behavior in rats. In W. K. Honig & G. Fetterman (Eds.), *Cognitive aspects of stimulus control* (pp. 385-404). Hillsdale, NJ: Erlbaum.
- CAPALDI, E. J. (1994). The sequential view: From rapidly fading stimulus traces to the organization of memory and the abstract concept of number. *Psychonomic Bulletin & Review*, **1**, 156-181.
- CAPALDI, E. J., MILLER, D. J., ALPTEKIN, S., & BARRY, K. (1990). Organized responding in instrumental learning: Chunks and superchunks. *Learning & Motivation*, **21**, 415-433.
- CAPALDI, E. J., NAWROCKI, T. M., MILLER, D. J., & VERRY, D. R. (1986). Time between events as a retrieval cue: Recall and the temporal similarity between the storage and retrieval intervals. *Journal of Experimental Psychology: Animal Behavioral Processes*, **12**, 258-269.
- CAPALDI, E. J., NAWROCKI, T. M., & VERRY, D. R. (1983). The nature of anticipation: An inter- and intraevent process. *Animal Learning & Behavior*, **11**, 193-198.
- CAPALDI, E. J., & VERRY, D. R. (1981). Serial order anticipation learning in rats: Memory of multiple hedonic events and their order. *Animal Learning & Behavior*, **9**, 441-453.
- CAPALDI, E. J., VERRY, D. R., NAWROCKI, T. M., & MILLER, D. J. (1984). Serial learning, interitem associations, phrasing cues, interference, overshadowing, chunking, memory, and extinction. *Animal Learning & Behavior*, **12**, 7-20.
- EBBINGHAUS, H. (1885). *Memory: A contribution to experimental psychology* (H. A. Ruger & C. E. Bussenius, Trans.). New York: Dover.
- FOUNTAIN, S. B., HENNE, D. R., & HULSE, S. H. (1984). Phrasing cues and hierarchical organization in serial pattern learning by rats. *Journal of Experimental Psychology: Animal Behavior Processes*, **10**, 30-45.
- HAGGBLOOM, S. J. (1981). Blocking in successive differential conditioning: Prior acquisition of control by internal cues blocks the acquisition of control by brightness. *Learning & Motivation*, **12**, 485-508.
- HAGGBLOOM, S. J. (1983). Blocking and partial reinforcement: Effects of N-R transitions early vs. late in training on resistance to extinction. *Bulletin of the Psychonomic Society*, **21**, 153-156.
- HULL, C. L. (1943). *Principles of behavior*. New York: Appleton-Century-Crofts.
- MACUDA, T., & ROBERTS, W. A. (1995). Further evidence for hierarchical chunking in rat spatial memory. *Journal of Experimental Psychology: Animal Behavior Processes*, **21**, 20-32.
- NEELY, J. H., & WAGNER, A. R. (1974). Attenuation of blocking with shifts in reward: The involvement of schedule-generated contextual cues. *Journal of Experimental Psychology*, **102**, 751-763.
- TERRACE, H. S. (in press). The comparative psychology of serially organized behavior. In S. B. Fountain, M. D. Bunsey, J. H. Danks, & M. K. McBeath (Eds.), *Sequential and temporal organization: Behavioral and biological perspectives*. Thousand Oaks, CA: Sage.
- THORNDIKE, E. L. (1911). *Animal intelligence: Experimental studies*. New York: Macmillan.
- WICKELGREN, W. A. (1964). Size of rehearsal group and short-term memory. *Journal of Experimental Psychology*, **68**, 413-419.

(Manuscript received August 3, 1998;
revision accepted for publication November 18, 1998.)

Nominations for the Editorship of *Psychobiology*

Nominations are solicited for the editorship of *Psychobiology*. The term of the present editor, Raymond Kesner, expires at the end of 2000. The new editor will begin an official 4- or 5-year term on January 1, 2001, and will begin to receive manuscripts early in April 2000. The Publications Committee of the Psychonomic Society expects to appoint the new editor by December 1999.

Nominations (including self-nominations) should be submitted by July 1, 1999, to:

Michael I. Posner
Chair, *Psychobiology* Search Committee
Department of Psychiatry, Box 171
Cornell Medical College
525 E. 68th St.
New York, NY 10021