The Psychological Record, 1999, 49, 299-325

DECREASE OF PRECURRENT BEHAVIOR AS TRAINING INCREASES: EFFECTS OF TASK COMPLEXITY

JORGE M. OLIVEIRA-CASTRO, DOMINGOS S. COELHO,
and GARDÊNIA A. OLIVEIRA-CASTRO
Universidade de Brasília

When someone is described as memorizing ~ phone number, part of what is being asserted is that the person is capable of dialing the number without looking it up in the directory. Such responses, which may decrease and stop occurring as training increases, can be interpreted as nonrequired precurrent behavior. In different experiments, participants could look up an auxiliary screen to see the numbers (Experiment 1) or arbitrary characters (Experiment 3) corresponding to different shapes. In Experiment 2, a typing task with a covered keyboard was used, in which participants could look up an auxiliary screen to see key positions. Duration of precurrent response, divided by correct current responses, decreased as a linear function of the logarithm of trials in all three experiments. In Experiment 3, the complexity of the task was changed, by altering the number of responses to be learned per pair, per position, and in the total task. Results indicated that these variables produced systematic effects on performance and are compatible with an interpretation of task complexity based upon the quantification of the programmed contingencies of reinforcement.

In ordinary language, people are often described as doing things in the head. When someone makes mental calculations, we say that the person solved the problems in his or her head, or mentally, and the

---

Portions of Experiments 1 and 2 were presented at the 21st and 23rd Annual Meeting of the Sociedade Brasileira de Psicologia, October 1991 and 1993, Ribeirão Preto, Brazil, and at the 21st Annual Convention of the Association for Behavior Analysis, May 1995, Washington DC. Portions of the data of these experiments appeared in a Portuguese publication (see Oliveira-Castro, 1993). Experiment 3 was part of Domingos S. Coelho master's thesis, presented at the Universidade de Brasilia in November 1995. Portions of Experiment 3 were presented at the 25th Annual Meeting of the Sociedade Brasileira de Psicologia, October 1995, Ribeirão Preto, Brazil, and at the 23rd Annual Convention of the Association for Behavior Analysis, May 1997, Chicago. The authors thank Ana Lídia G. Gama, Eileen P. Flores, Glaycilene S. Oliveira, Pedro Oliveira, Carlos B. A. Souza, and Eduardo Schwarz for help in the collection of data, and Dudley Terrell for helpful comments on previous versions of this manuscript. Financial support was received from CNPq, Brazil. Correspondence concerning this article should be addressed to Jorge M. Oliveira-Castro, Departamento de Processos Psicológicos Básicos, Instituto de Psicologia, Universidade de Brasilia, Campus Universitário Darcy Ribeiro, Brasilia, DF 70910-900, Brasil. Electronic mail may be sent via Internet to jocastro@unb.br.

numbers are said to have been added or multiplied in his or her head. After I sing a particular tune for a while, the melody may keep running through my head. Analyzing the use of the expression *in* my *head* in ordinary language, Ryle (1949) emphasized two important characteristics of such use. First, the concept has an undeniably metaphorical function. When someone makes mental calculations or has a melody playing in his or her head, no one would expect the numbers to appear on an x-ray of his or her brain or to hear a muffled melody by applying a stethoscope to his or her cranium.

Secondly, the concept has an indispensable negative function, which Ryle (1949) illustrated through an example. When the wheel-noises of the train make a melody run in my head, the wheel-noises are audible to the other passengers, but the melody is not (p. 36). To assert that I have a melody running through my head is a way of denying that there is really any music being played by an orchestra or a record player. This negative function of the use of *in* my *head* indicates that when we say, in ordinary language, that the boy solved the arithmetic problems in his head or mentally, we are asserting that the boy solved the problems without writing down or looking at the numbers on paper or blackboard; neither has he spoken nor heard the numbers. One of the main functions of the expression *in the head,* in this context, is to indicate that some things did not occur. By the same token, when someone is described as keeping a phone number in his or her head, part of what is being said is that the person is capable of saying, writing down, or dialing the number, without looking it up in the directory or asking someone else. Although looking the number up or asking someone may have been a necessary condition for correct dialing at some point during the learning process, this type of behavior ceased to be necessary and stopped occurring. Examples of this kind of behavior can be identified in almost any task, such as looking at the pedals when learning to drive a car, or looking at the keyboard when learning to type, or listening to the teacher when learning to pronounce a new foreign word, or looking at a multiplication table when solving arithmetic problems. In all these cases, correct responding may occur, after some training, without the emission of such responses, which may drop out from the original response sequence (Oliveira-Castro, 1992, 1993).

In operant terms, this type of behavior may be interpreted as a kind of precurrent (or mediating) behavior, a concept used by Skinner (1953, 1957, 1968, 1969) to refer to responses that increase the likelihood of other response (current) occurring or being reinforced. These responses, such as looking up a phone number in the directory, may increase, at least at the beginning of training, the likelihood of other (current) responses being reinforced, such as dialing the correct number. Considering, moreover, that precurrent contingencies may differ with respect to several characteristics, as suggested by Polson and Parsons (1994), these responses may be described as signaled, for they produce stimulus changes (e.g., the number in the directory) correlated with changes in the reinforcement parameters for the current response (e.g., dialing the number) and are not required by the programmed contingencies (e.g., dialing could be reinforced without looking up the number). This type of behavior could be then characterized as signaled and nonrequired precurrent responses that may, with increased training, decrease and stop occurring without disrupting current responding.

Experimental investigations of precurrent behavior have been, for the most part, concerned with the influence of precurrent responses on current responding in temporally defined reinforcement schedules (cf. Parsons, Taylor, & Joyce, 1981). Collateral responses with precurrent functions have been described in *differential-reinforcement-of-low-rate* schedules performed by humans (e.g., Bruner & Revusky, 1961), monkeys (e.g., Hodos, Ross, & Brady, 1962), and rats (e.g., Laties, Weiss, Clark, & Reynolds, 1965; Laties, Weiss, & Weiss, 1969). The influence of precurrent responses on performance in delayed-choice procedures, such as delayed-matching-to-sample and alternation, has been reported in experiments with humans (e.g., Parsons et al., 1981; Torgrud & Holborn, 1989) and pigeons (e.g., Blough, 1959; Eckerman, 1970; Hearst, 1962; Jans & Catania, 1980; Shimp & Moffitt, 1977). In all the experiments in which precurrent responses were prevented, prohibited, or disrupted after conditions in which they occurred, this type of manipulation produced disruption of current responding. Such manipulations decreased interresponse time in differential-reinforcement-of-low-rate schedules, reducing obtained reinforcement rates (cf. Hodos et al., 1962; Laties et al., 1965, 1969), and accuracy of choice with increasing delay in delayed-choice procedures (cf. Blough, 1959; Jans & Catania, 1980; Parsons et al., 1981; Torgrud & Holborn, 1989). These results suggest that there are important differences between the precurrent contingencies in effect in such experiments and those in effect in situations in which responses similar to looking-up-the-phone-number-in-the-directory occur, for these latter may, apparently, stop occurring without disrupting current responding.

Responses similar to looking-up-the-phone-number-in-the-directory resemble those responses that have been called *observing behavior.* This expression has been widely used to refer to (a) responses that generate discriminative stimuli otherwise not present in the situation, such as a response that produces the stimuli associated with each component of mixed schedules of reinforcement (cf. Catania, 1992, p. 174; l?msmoor, 1983); and (b) responses to the sample in a typical matching-to-sample procedure (cf. Catania, 1992, p. 151). In both cases, such responses could be interpreted as precurrent, considering that they increase the likelihood of more efficient response patterns (in the case of *a)* or are required for reinforcement of current responses (in the case of *b)* as in a typical response chain (cf. Polson & Parsons, 1994). Responses such as looking-up-the-phone-number-in-the-directory would therefore be functionally similar to observing responses. Despite this general similarity, research on observing behavior has been primarily concerned with testing different hypotheses (conditioned reinforcement or information) related to the variables that maintain observing responses rather than identifying the variables responsible for their decrease. As a matter of fact, in those few experiments in which the frequency of observing responses decreased with increasing training, this result has been regarded as unexpected and difficult to explain (e.g., Bickel, Higgins, & Hughes, 1991; D'Amato, Etkin, & Fazzari, 1968; Mueller & Dinsmoor, 1984; Ohta, 1987). The experimental situations used to investigate observing behavior seem also to differ from the situations in which responses such as looking-up-in-the-directory occur.

Polson, Grabavac, and Parsons (1997) conducted a series of experiments to investigate transfer effects of reversing intraverbal responses in a situation similar to the phone number example. On each trial of the adopted procedure, an English (or French) word was presented on a computer screen and participants had to type its corresponding French (or English) word. Participants could type the word, skip the trial, or get a hint, in which case each key press would show one of the letters, in the correct sequence, of the word to be typed and the typing response would be counted as incorrect. This setting a hint was somewhat analogous to the looking-up-in-the-directory response. Data from several participants showed perfect response accuracy in some conditions, indicating that the precurrent response (getting a hint) stopped occurring without disrupting the current response (word typing). When compared to previously described results, these results raise several questions. Why, in this experiment, precurrent responses decreased and stopped occurring without disrupting current responding, and, in temporally defined schedules, current responses were disrupted when they were prevented? Why precurrent responses would be expected to stop occurring in the phone number example, whereas they were not expected to decrease in typical procedures used to investigate observing responses?

These discrepant results may perhaps be explained by examining the possibility of transfer of stimulus function in the experimental situations. The decrease and eventual nonoccurrence of the looking-up-in-the-directory response, as well as the response of getting a hint, could be explained as a case of transfer of stimulus function. Being presented, at the beginning of training with a given name, the person looks it up in the directory, sees the phone number corresponding to that name, and dials it, producing thus the consequences for correct responses (such as talking to the person). As this procedure is repeated several times, that is, as that particular name-number pair is repeated, the stimulus function of the number, which influences what the person dials, is transferred to the name. After some training the person can then dial the correct number without looking it up in the directory, for the name at that point also influences what the person dials. This would explain why these responses may decrease and stop occurring without disrupting current responding.

Moreover, in situations similar to the phone number example there is a high correlation, close to one, between the stimuli that set the occasion for the current response, such as the name of whom one needs to call, and the stimulus produced by the precurrent response, such as the number one sees in the directory. As persons' names and phone numbers are, usually, consistently associated, the correlation between them would be high. If the pairing of names and phone numbers changed every time one dialed (e.g., Mary's number would become John's, John's would become Phillip's, and so on), the correlation between names and phone numbers would be low, the situation would prevent transfer of stimulus function, and we would have to look up the number (not in the directory but in some on-line system) every time we called someone. The situation would then be analogous to typical procedures used to investigate observing responses, which minimizes or prevents transfer of stimulus function. This is the case when mixed schedules of reinforcement may be transformed into multiple schedules with the emission of an observing response (the case of a above). As the components of mixed schedules alternate randomly, without any signal, the possibility of transfer of stimulus function is minimal, that is, without the emission of observing responses there is a much reduced possibility of the occurrence of the most efficient response pattern associated with each component. In the case of responses to the sample stimulus in matching-to-sample procedures (the case of *b* above), any transfer of stimulus function is prevented, for the subject could not possibly choose, above chance level, the correct comparison stimulus without seeing (if visual discriminations are required, although it could be hearing, tasting, and the like) the sample.

Responses such as looking-up-the-phone-number-in-the-directory could therefore be interpreted as nonrequired and signaled precurrent responses occurring in situations in which there is possibility of transfer of stimulus function. Although the results obtained by Polson et al. (1997) may be cited as experimental demonstrations of precurrent responses that stopped occurring without disrupting current responding, they do not allow a separate analysis of the decrease of precurrent responses. The measure adopted in the experiment, rates of correct responding, did not separate the duration of precurrent responses (i.e., getting hints) from the latency and duration of current responses (i.e., word typing). The main purposes of the present experiments were to develop laboratory tasks that would allow an experimental analysis of this type of precurrent behavior and to investigate the effects of variables related to task complexity on their decrease. In Experiment 1, a task analogous to the phone number example was used, in which, upon being presented with one of eight different shapes (analogous to person names), participants had to type the number (e.g., 53481) corresponding to it. On each trial, participants could look up the number corresponding to the presented shape by activating an auxiliary screen (precurrent response). In Experiment 2, a typing task was adopted, in which participants should type, using a covered keyboard, the key corresponding to a letter or character presented on the screen. On each trial, participants could look up a drawing containing ali key positions by activating an auxiliary screen (precurrent response). In Experiment 3 a task similar to that used in Experiment 1 was adopted, in which participants had to learn pairs of shapes and sets of arbitrary characters. The complexity of this task was manipulated in different experimental conditions.

Experiment 1

In .Experiment 1, a paired-associates task was used in which, given an arbitrary shape presented on a computer screen, participants had to type its corresponding number. Participants could look up each shape's corresponding number by activating an auxiliary screen. The precurrent response was thus looking up the auxiliary screen and participants' task was to memorize the numbers.

Method

Partícípants. Twenty-one persons (9 men and 12 women), ranging in age from 18 years to 42 years, volunteered to participate in the experiment; most (16) were students at the Universidade de Brasília and none had participated in a similar experiment conducted as part of a psyc.hology course laboratory practice. Points obtained during the session were exchanged for money. The highest value paid per session was approximately $5.00, although the exact value was difficult to calculate because of the high inflation during the period.

~quípm,nt. One personal computer (Swan 386 SX) with color rnonitor and keyboard, located in a room without sound attenuation, was used. Computer programs were written in Pascal (version 5.5).

Procedure. Each participant read the following instructions, written in Portuguese, before starting the session:

> Thank you for your participation in this experiment about memory. Vour task is to memorize some numbers. Each number is associated with a symbol.. Read carefully the following instructions.
>
> A symbol will appear on the screen. PRESS the up-arrow key to see the number corresponding to the symbol, which will appear on a white screen. PRESS the down-arrow key to return to the blue screen, when you are ready to write the number. In order to write the number, TYPE one digit at a time, using the keys located on the right-hand side of the keyboard. Attention! Vou will not be able to correct your responses. If you type any incorrect digit, continue. Type the next digit that makes up the number. Vou may look up the white screen whenever you find it necessary, but remember that your task is to MEMORIZE the numbers. If you type ali the correct digits, the computer will produce a high-pitch tone. If any digit is incorrect, ~he computer will produce a low-pitch tone. Vou may look up the white screen whenever you find it necessary. However, every time you type a number correctly without looking up the white screen, you will score a point on the marker located on the left of the screen. Each point you score will be exchanged for Cr$ 20,00. After the tone presentation, a new symbol will be presented, and you should follow the same steps described above. Press the [Enter] key in order to see some examples.

Two shape-number pairs were presented for two trials each as examples, which were followed by the experimenter, who explained' the procedure. After these examples, the participant had to press the [Enter] key to start the session and the experimenter left the room.

The sequence of events during a trial was: (a) an arbitrary shape was. ~resented on the top central part of a blue screen; (b) if the participant pressed the up-arrow key, the word "Wait," written in white color, appeared on the top central part of the blue screen for a short period of time *(t1); (~)* after *t1,* a five-digit number appeared on the top central part of a white screen; (d) when the participant pressed the down-arrow key, the word "Wait," written in blue color, appeared on the center of t~e white screen for 0.5 s; (e) after this, the blue screen app~~red without the shape, with the cursor blinking at its center; (f) the participant typed the number, using the number keys located on the left-hand side of the keyboard, which appeared on the center of the blue screen; (g) aft~r 0.5 s, a high-pitch tone (800Hz - correct) or a low-pitch tone (1OO~z.-incorrect) was p~esented depending on the number typed by the partícipant: (h) after which, a medium-pitch tone (450Hz) and the word "Interval," written in red, were presented for 0.5 s. After this, another arbitrary shape was presented, starting thus a new trial.. Steps *b* thro.ugh e .(pr~curr~nt response) could be skipped by the participant on any qrven trial, In whlch case the first typing response erased the shape from the blue screen. If the participant typed a correct number without looking up the auxiliary screen (skipping steps *b* through e) a point was added in the marker shown on the top left-hand comer of the blue screen. As informed in the instructions, participants could not correct their response, for the only active keys were the number keys, up- and down-arrow keys, and the [Enter] key.

Eight shape-number pairs were used. The shapes were similar to those used in Experiment 3 (see Figure 3) and the numbers were: 53481,72935,86593,35867,94218,46359,68742,29674. These numbers were created following three rules: (a) digits should be different than .z~ro, (b) no digit should be repeated in the same number, and (c) no diqlt should be repeated in the same position. The value of *t1* was equal to 0:5, 1:0, 2.5, and 5.0 s for each two of the eight shape-number parrs, but Identlc~1 for ali participants. Considering that this manipulation had no systernatic effects upon the results, it will be disregarded in the pres.ent paper. The sessi~n ended after 24 blocks of trials during which ali eiqht shape-number parrs were randomly presented once (Le., total of 192 trials).

The frequency and duration of precurrent responses, that is, activating the white screen containing the number, and the. number of digits typed correctly were recorded for each trial of each participant,

Results and Discussion

The frequency of precurrent responses as a function of blocks of trials showed that, for most participants, the frequency was equal to 8.0 on the first trials (one precurrent response for each pair), decreasing to zero or close to it by the 24th trial block. Considering that the procedure did not impose any restriction on the frequency or duration of precurrent responses, that is, participants could look up the white screen whenever and for as long as they wanted, and that the sum of durations would include the frequency, but not the reverse, the duration of precurrent response, rather than its frequency, seems to be a more interesting measure. This would have been different if the duration of precurrent response had been fixed at a short value. On balance, the duration of precurrent response, by itself, does not seem to be an adequate measure of learning, for individuals could stop looking up the white screen, despite making errors. For this reason, the main results will be described in terms of duration of precurrent responses divided by the number of correct current responses (digits typed).

Figure 1 shows the duration (seconds) of precurrent responses divided by the number of correct current responses as a function. of trial blocks for each participant in Experiment 1. Each point was obtained by the sum of durations of precurrent responses divided by the number of correct digits typed, for all eight shape-number pairs, on each trial block. The cases where errors occurred and the sum of precurrent response duration was equal to zero were treated as missing points.

Precurrent response duration per correct decreased, as a negatively accelerated curve, as the number of trials increased for all participants. Despite this similar decreasing pattern, individual differences in performance were observed, as indicated by the value of precurrent response duration per correct on the 24th trial block and its maximum value in the session (note the differences in graph scales). Whereas precurrent response duration per correct was equal to zero for Participant 14 on the 17th trial block, this same measure was equal to 0.59 s for Participant 26 on the last trial block. The maximum value of precurrent response duration per correct, on any given trial block, varied across individuals from 0.82 s (P17) to 7.83 s (P27). Another aspect that calls attention in the figure is the fact that precurrent response duration per correct showed, for three participants (P7, P10, P14), values larger than zero after trial blocks in which it was equal to zero, indicating that precurrent responses did not necessarily stop occurring in an all-or-none fashion. These results corroborate those obtained by Polson et al. (1997) in showing that precurrent responses may decrease, and even stop occurring, without disrupting current responding.

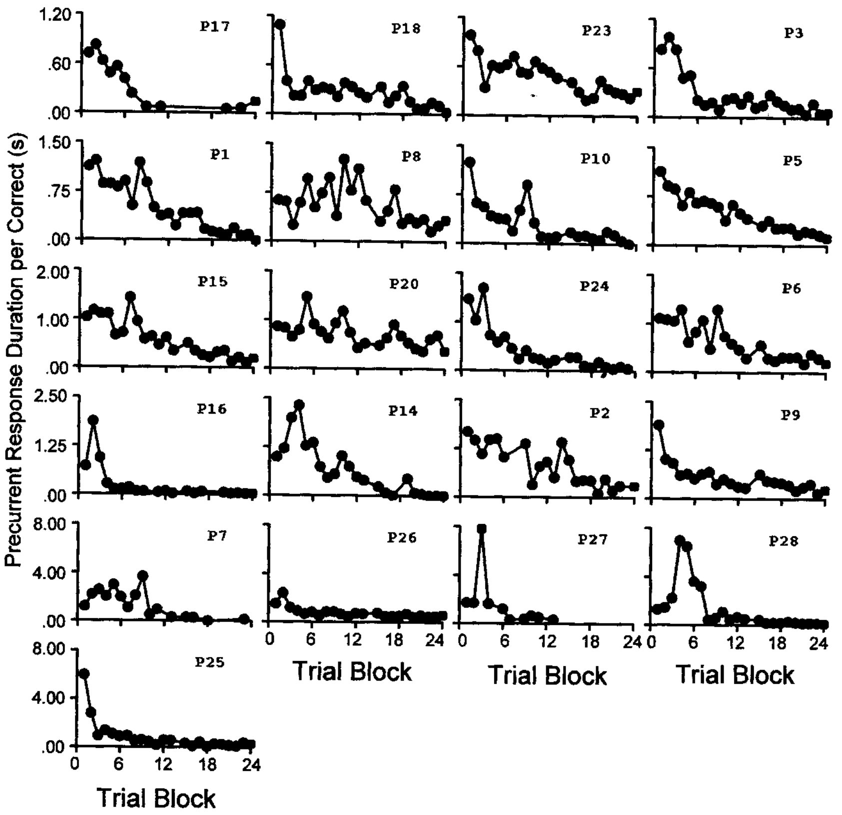

Figure 1. Duration of precurrent response (s) divided by the number of correct current responses as a function of trial blocks, calculated for each participant in Experiment 1.

Experiment 2

Considering, as mentioned previously, that this type of precurrent behavior may be found in several tasks, Experiment 2 was conducted with the purpose of investigating such behavior in a different task. A typing task was used, in which the keys on the keyboard were covered and the participants could look up the keyboard characters on an auxiliary screen. The task also differed from that used in Experiment 1 with respect to two other things: (a) All the characters to be learned (or memorized) were presented together on the auxiliary screen, whereas in the previous experiment each number was presented individually; and (b) typing is a familiar task, with which participants had different levels of experience.

Method

Participants. Thirteen individuals (6 men and 7 women), ranging in age from 19 to 35 years, students or *staft* ~embers at the Universidade de Brasília, who asserted that they did not know how to type, volunteered to participate in the experiment.

Equipment. One personal computer (Swan 386-SX) ~Ith color monitor and a keyboard covered with black tape, located In a r?om without sound attenuation, was used. Computer programs were wntten in Pascal (version 5.5).

Procedure. Participants read instructions analogous to those u.sed In Experiment 1 with the additional req~est of pr~~sing each key. wlth the correct finger, which were indicated m the aux!hary scree~. Tnals were analogous to those used in Experiment 1. A tnal began with a letter or character presented at the top central part of a blu~ screen (i.nstead of a shape). The auxiliary screen (analogous to the whlte s~reen m the other experiment) contained a drawing of the keyb?ard showinq ali letters and characters used in the experirnent, on their respective keys, and the drawing 0l two hands, left and right, on th~ir respective sides of the keyboard. The hand fingers were drawn In the same color as the letters/characters they should press (a total of eight colors).

A total of 46 letters and characters were used. These included ali capital letters (26), ali digits (10), and ten characters ({-}, {=}'. {I}, {:}, q, {.}, {?}, {"}, {r}, {]}). With the exception of the [Enter] key whlch rernained uncovered ali keys were covered with black tape on the keyboard. Ali k~ys not used i~ the procedure appeared empty on the auxiliary screen drawinq. Only one character/number was chosen for those keys that contam, m standard keyboards, two characters or a character and a number.

Ali 46 letters/characters were randomly presented once dunng each trial bloc~ and the session ended after 30 trial blocks. There was no record.oí or control over the fingers participants used to press the keys. The time between an [Enter] key press and the presentation of the auxiliary screen (analogous to *t1* in Experiment 1) was equal to 0.5 s.

Results and Discussion

Figure 2 shows precurrent respo.n~e du~ation pe~ correct as a function of trial blocks for each particlpant In Expenment 2. Data calculations were analogous to those used in Experiment 1. T~o participants interrupted the session on the 15th (P11) and 19~h (P5) tnal blocks, saying that they had other commitments (some sessions lasted for alrnost 2 hours), whereas three other participants, despite the fact that they still made errors, stopped looking up the auxiliar.y screen on the 14th (P7), 15th (P8), and 18th (P6) trial blocks. The available data f~om these individuais are presented in the figure, whereas the other potnts were treated as missing.

Precurrent response duration per correct decreased as tnal blocks increased for ali participants. As in Experiment 1, such decreases were negatively accelerated. Individual difterences in perform~nce were also observed. For Participant 9, precurrent response duratíbn per correct was zero on the third trial block, whereas for six participants (4, 11, 8, 7, 6, 5), including the two that ínterrupted the session, this measure was still larger than zero by the end of the sessíon. The maximum value of precurrent response duration per correct, on any gíven trial block, varied across participants from 1.63 s (P9) to 26.59 s (P14). These wide individual difterences may have been caused by participants' different leveis of experience with typing tasks. As in Experiment 1, precurrent response duration per correct larger than zero after trial blocks in which it was equal to zero were also observed. The task used in Experiment 2, which included 46 one-digit pairs, seemed to be more complex than those used in Experíment 1, whích included 8 five-digit pairs, as can be seen by the higher values of precurrent response per correct observed in the last experíment.

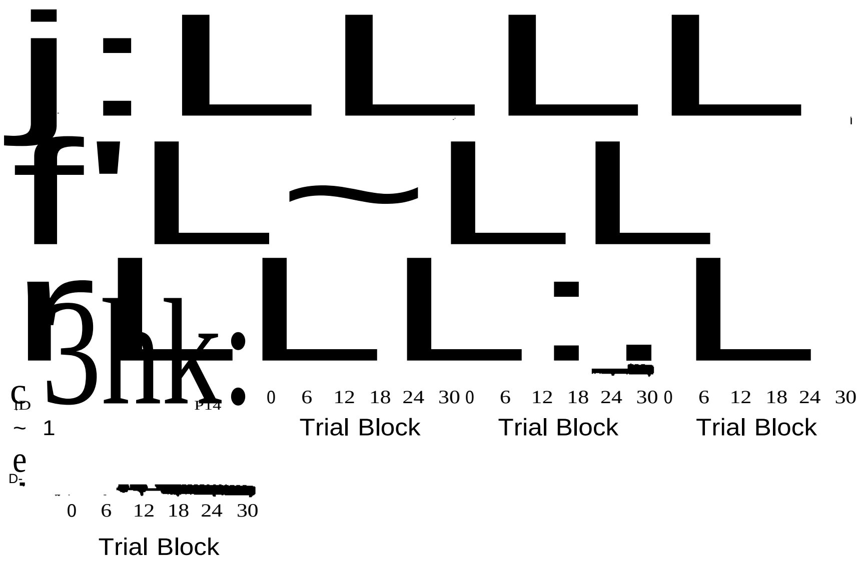

Figure 2. Duration of precurrent response (5) divided by the number of correct current responses as a function of trial blocks, calculated for each participant in Experiment 2.

Experiment 3

The complexity of the task is among the variables that may influence the decrease of this type of precurrent behavior, for one would expect changes ín the duration, or even in the functíon, of precurrent behavior as task complexity changes. In some highly complex tasks, precurrent behavior may never stop occurring (e.g., memorizing the phone books of New York City), whereas in simpie tasks (e.g., memorizing one phone ~umber) this behavior may stop occurring after one trial, in which case ItS decrease would not be negatively accelerated (cf. Skinner, 1953, p. 61). Considering that, in most examples of human behavior, higher leveis of performance are formed from previous leveis in which this type of precurrent responses used to occur and stopped occurring (e.g., most adults can multiply large numbers without looking up multiplication tables), an analysis of the effects of task complexity on this type of precurrent behavior may be relevant to theoretical interpretations of complex human behavior and the development of teaching technologies.

An analysis of the literature concerned with the effects of task complexity, however, indicates that there is no general and widely accepted description, theory, or analysis of task complexity in psychology. Attempts to describe task complexity have included one or more of the following items: (a) task characteristics, such as type of instructions or stimuli sensory modality, (b) the necessary behavior for certain level of performance, (c) the actual behavior emitted by the individual, or (d) the skills necessary for good performance (cf. Hackman, 1969; Wood, 1986).

Although the level of performance of any individual in any task is, undoubtedly, the product of the interaction among certain characteristics of the task and the initial repertoire (i.e., skills or previous training) of that individual, an analysis of the logic of the concept of task complexity suggests that the concept is related to task characteristics and not to individuals' skills, as can be seen by the following argument. As an individual practices a given task, although his or her performance might improve on that task and the individual may be described as acquiring or improving some skill, it would make little sense to assert that the complexity of the task decreases as practice increases. Piloting a modern fighter, or writing, does not become less complex because some people can do it well and with ease. Despite the fact that in some contexts the concepts of complexity and difficulty are used as synonyms, they show an interesting asymmetry of use. Whereas *difficult* implies necessarily lower performance levels (when compared to easy, across tasks or individuals), the same does not happen with *complex;* a complex task may be very easy for a given individual or after some training, as exemplified earlier. According to this analysis, therefore, individuals' repertoire should always be considered when referring to task difficulty but not necessarily to the notion of task complexity, the latter being related to characteristics of the task independently of individuals' repertoire. Based upon different arguments from the one presented here, Hackman (1969) and Wood (1986) reached a similar conclusion when defending that descriptions of task complexity should take into account task characteristics and required behavior, disregarding individuals' skills and the actual behavior emitted in the situation. If one accepts this argument, however, there would still remain the question concerning the kind of task characteristics that should be included in the description of task complexity.

Considering that any task specifies what responses will be correct (or reinforced or successful) in what situations (discriminative stimuli), one possible way of analyzing task complexity, which would be compatible with the theoretical considerations made so far and with an

operant interpretation of task complexity, would be to interpret any task as a set of programmed contingencies of reinforcement. The quantification of such contingencies would then provide a measure of task complexity. This approach would avoid the difficulty of selecting the dimensions that should be described, considering that a complete description of task characteristics is not possible, as pointed out by some authors (cf. Hac~man, 1969; ~ood, 1986). This type of analysis would describe those dimensions which are functionally specified by the programmed contingencies of reinforcement.

In the task used in Experiment 1, for example, the contingencies specified that a given response (typing one character) would be correct only if it occurred in the presence of the correct shape and in the correct position. Shape and position were then two functional dimensions in that task because the consequence delivered after each response ("Right!" or "Wrong!") could change depending on them.

In order to test the adequacy of this type of analysis for the investigation of the effects of task complexity on precurrent behavior it would be convenient to manipulate a variable whose effects have already been described. Despite the lack of consensus or agreement concerning the best way of theorizing about task complexity, the effects of one variable has been systematically mentioned in investigations following different theoretical traditions. Increases in the number of responses to be learned (or that are possible) in a given situation have been associated with decreases in performance in several different experimental contexts: (a) The increase in the number of items to be memorized in serial learning tasks increased the time to learn each item (cf. Eb~ingha.us, 188~/1964; for opposite results, see Deese, 1958); (b) according to information theory, the number of response alternatives or possible stimuli in the situation, determines the complexity of different tasks (e.g., Coren & Ward, 1989; Simon, 1972, 1974); (c) decision making research has suggested that the number of choice alternatives is one of the variables that influences task complexity (e.g., Brehmer, 1992; Kerst~olt, 1992; Payne, 1982; Sündstrom, 1987); (d) the cued recall of an item decreases as the number of items associated with the same cue increases (e.g., Bãuml, 1997; Ratcliff, Clark, & Shiffrin, 1990); and (e) the increase in the number of responses associated with a given stimulus in paired-associates tasks reduced performance in transfer tasks, when compared to the increase in the number of stimuli associated to the same responses (e.g., Postman, 1972).

Considering, despite differences in the theoretical approaches this wide range of investigations that have reported effects on performance of the number of responses to be learned (or that are possible) in a given task, the present experiment investigated the effects of this variable on precurrent behavior. The procedure used was identical to Experiment 1, with the difference that instead of using shape-number pairs the present experiment used pairs of shapes and five-digit sets of arbitrary characters, which should reduce the influence of participants' previous experiences upon the results. The first session was identical for ali participants, which makes possible a comparison of the decrease of precurrent behavior with arbitrary materiais with the results obtained in previous experiments. In different experimental conditions, the number of responses to be learned was manipulated, separately, per shape (number of different arbitrary characters that formed each five-digit set), per position (number of different arbitrary characters occurring in each of the five positions across pairs), and in the total set of material (number of different arbitrary characters that formed ali pairs and positions). These increases in the number of different responses to be learned could be interpreted, according to the analysis of task complexity proposed here, as decreases in the programmed probability of reinforcement given each discriminative dimension (shape and position) and in the average programmed frequency of reinforcement in the task.

Method

Participants. Twenty-four Universidade de Brasília students (8 men and 16 women), ranging in age from 17 to 30 years, volunteered to participate in the experiment. Twenty participants received course credits that were proportional to the duration of their participation (0.5% of final grade per hour, up to 10 hours). The other four students volunteered to participate despite the fact that none of their courses would accept research participation credits.

Equipment and material. One personal computer (Swan 386 SX) with color monitor and a modified keyboard was used. The keyboard contained 46 characters chosen from the ASCII Table (178 through 223) and had all contrai keys covered with the exception of the up- and down-arrow keys and the [Enter] key. Figure 3 shows the pairs of shapes and sets of characters used. Each character that formed the five-digit sets of characters (second members of the pairs) was selected from one out of eight, arbitrarily defined, regions of the keyboard. Each arbitrary shape (first members of the pairs) was formed by repeating six times the same character. The equipment was located in a room without sound attenuation and computer programs were written in Pascal (version 5.5).

Procedure. Before starting the session, each participant read instructions similar to those presented in Experiment 1, the main difference of which was that participants were instructed to type codes instead of numbers. The sequence of events on each trial was identical to that in Experiment 1, with the time between an up-arrow key press and the presentation of the auxiliary screen (*t1* in Experiment 1) equal to 0.5 s. Participants typed the characters using the keys on the keyboard which contained the arbitrary characters over the normal letters.

Ali participants were submitted to five sessions, separated by a period of time ranging from 24 to 72 hours, each one of which being a different experimental condition. Condition 1 differed in several aspects from the other conditions: It was the first session for ali participants, eight shape-characters pairs were used (see Figure 3), the session ended when the participant had obtained 16 consecutive points in two consecutive trial blocks or after 24 trial blocks, the second member of each pair was formed by one character repeated five times, and it differed from ali other conditions with respect to more than one of the variables manipulated (see conditions below). Considering these differences and that the main purposes of this first session were to replicate previous experiments using arbitrary characters instead of numbers or letters and to make participants more familiar with the task, data from it will not be compared with those from the other four experimental conditions.

In Conditions 2 through 5, four shape-characters pairs were used (see Figure 3) and sessions ended when the participant had obtained 16 consecutive points in 4 consecutive trial blocks or after 48 trial blocks. These four conditions differed from each other on the basis of one or

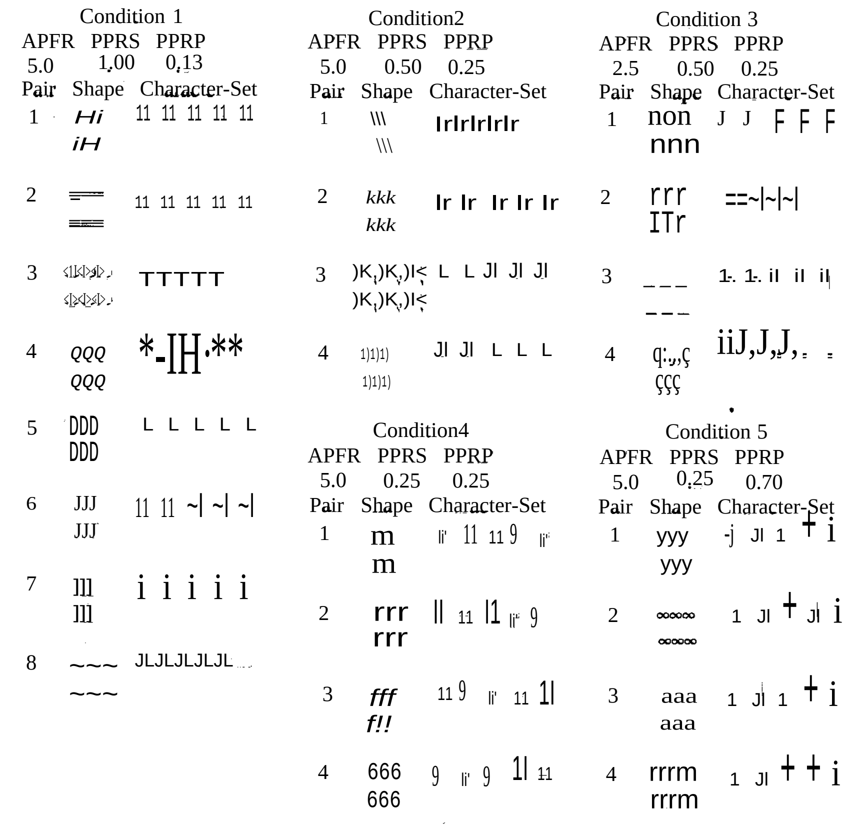

Figure 3. Pairs of shapes and arbitrary characters used in each condition in Experiment 3. The val.~es of a~erage programmed frequency of reinforcement (APFR), programmed pr.obabllIty of ~elnforce~~nt given a shape (PPRS), and programmed probability of relhforcement glven a posltlon (PPRP) are indicated.

more of the following variables: average number of different responses per pair, average number of different responses per position, and number of different responses in the total set of pairs. Each different arbitrary character to be typed was counted as a different response, considering that a change in any one of them would change the consequence (Right or Wrong) for the response series. As mentioned previously, these manipulations could be interpreted in terms of quantified programmed contingencies of reinforcement. . In this type of task, a response (i.e., typing an arbitrary character) would be reinforced (Le., correct) only if it occurred in the presence of the appropriate shape and in the correct position. Shape and position would be two discriminative dimensions, that is, correct responding depends on them. Moreover, changes in the total number of different responses (Le., arbitrary characters) to be learned could be interpreted as changes in the average programmed frequency of reinforcement in the task. **In** other words, participants had to learn the responses and the appropriate occasions to emit them. These variables can be quantified in the following manner: (a) average programmed frequency of reinforcement: . obtained by dividing the number of possible correct responses in any given trial (40 in Condition 1 and 20 in Conditions 2 through 5) by the number of different responses (8 in Conditions 1 and 3, and 4 in the other conditions); (b) average programmed probability of reinforcement for any response given a shape: obtained by dividing the average frequency of reinforcement given a shape (5 in Condition 1, 2.5 in Conditions 2 and 3, and 1.25 in Conditions 4 and 5) by the number of possible correct responses given a shape (equal to the number of position}; 5 in ali conditions); (c) average programmed probability of reinforcement for any response given a position: obtained by dividing the averáge frequency of reinforcement given a position (2.8 in Condition 5 and 1 in ali other conditions) by the number of possible correct responses given a position (equal to the number of shapes; 8 in Condition 1 and 4 in ali other conditions).

The values of average programmed frequency of reinforcement, programmed probability of reinforcement given a shape, and programmed probability of reinforcement given a position were, respectively, the following: Condition 1 = 5.0, 1.00, 0.13; Condition 2 = 5.0, 0.50, 0.25; Condition 3 = 2.5, 0.50, 0.25; Condition 4 = 5.0, 0.25, 0.25; and Condition 5 = 5.0, 0.25, 0.70 (see Figure 3). Therefore, Conditions 2 and 3 differed only with respect to the average programmed frequency of reinforcement, Conditions 2 and 4 differed only with respect to the programmed probability of reinforcement given a shape, and Conditions 4 and 5 differed only with respect to programmed probability of reinforcement given a position.

The sequence of presentation of Conditions 2 through 5 and the character set used were partially balanced across four groups with six participants each. Character set refers to the specific characters used, independently of the condition, for example, four of the eight characters used to form the set for Condition 3 for Group 1 (see Figure 3) were used to form the set for Condition 2 for Group 3. The combination of ?o~dition~, s~ssions, an? character sets used for each group are indicated In Figure 5. Decimal numbers tor-character sets indicate that four of the eight characters of a given set were used (e.g., Character - Set 3.1 consisted of four characters which were chosen from the eight characters of Character - Set 3).

Results and Discussion

Session 1. Figure 4 shows duration (seconds) of precurrent responses divided by the number of correct current responses as a functi?n of trial blocks for each participant in Session 1 (Condition 1). Duration of precurrent response per correct decreased, as a negatively accelerated curve, as the number of trial bloéks increased for ali participan!s. Despite the fact that a similar decreasing trend was observed In the data from each participant, individual differences occurred with respect to the highest duration of response on any given

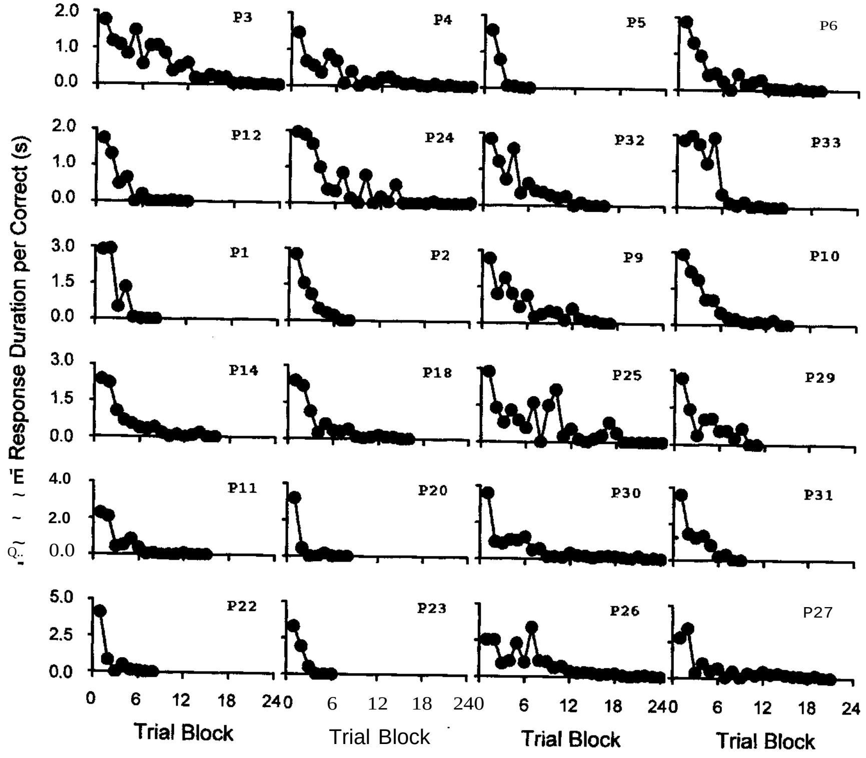

Figure 4. Ouration of precurrent response (s) divided by the number of correct current responses as a function of trial blocks, calculated for each partícipant in Sessíon 1 in Experíment 3.

trial block, which ranged from 1.46 s (P4) to 4.95 s (P22), and with respect to the number of trial blocks necessary to end the session, which ranged from 6 trial blocks for one participant (P5) to more than 24 trial blocks for six participants (P3, P4, P24, P25, P26, and P30) who did not reach the criterion during the session.

The results from Experiments 1, 2, and 3 (Session 1) raise several questions concerning the adequate measures to be adopted in describing the duration of precurrent behavior, if one is interested in comparing individuais' performance or the effects of independent variables. As seen in the experiments, precurrent response duration per correct varied across individuais with respect to both its maximum value across trials and the trial on which it was equal to zero, in such a way that one could find, for example, the following two extreme patterns of performance: (a) high precurrent response duration per correct for a few number of trial blocks before precurrent response duration was equal to zero (e.g., patterns similar to P22 and P23 in Experiment 3, see Figure 4), and (b) low precurrent response duration per correct for a large number of trial blocks before precurrent response duration was equal to zero (e.g., patterns similar to P6 and P32 in Experiment 3, see Figure 4; notice that the scales are different). How should these two response patterns be compared? Which performance could be described, in general, as better? How could one analyze the effect of an independent variable in arder to evaluate whether it facilitates or hinders the decrease of precurrent behavior? In addition to ali these questions, the results showed that precurrent response duration per correct may be greater than zero after a trial on which it was equal to zero. This raises the problem of identifying the trial block on which precurrent behavior stops occurring, which may have theoretical and practical relevance in the prediction of performance across different tasks.

In view of the observed regularity in the decrease of precurrent behavior and the questions raised above, an attempt was made to describe the relationship between precurrent behavior and trials according to the following equation:

$$\text{Precurrent Duration/Correct} = b - a\ (\text{log Trials}) \qquad (1)$$

in which precurrent response duration per correct would be a linear function of the logarithm of the number of trials, and *b* and a would be empirically derived parameters. It should be emphasized that the equation is proposed here with the sole purpose of concisely describing the data, containing thus free parameters, which would be used as measures of precurrent behavior. Equation 1 was chosen based upon the following criteria: (a) the negatively accelerated shape of the observed decreases in precurrent behavior: (b) the fact that the duration of precurrent behavior may be equal to zero, which eliminates theoretically, for example, a power function; (c) the simplicity of calculation of a linear function; (d) the simplicity of interpretation of the parameters, for *b,* the

intercept, would be the estimated value of precurrent response duration per correct on the first trial block, whereas a, the slope, would be the rate of decrease of precurrent response duration per correct across trials.

The mean value of the determination coefficient *(F)* obtained from Equation 1, calculated from the data obtained for each participant *(N* = 58) in Experiments 1,2, and 3 (Session 1) was equal to .68 *(SO* = .20), indicating that the equation fits the data reasonably well. According to the equation, the parameters a and *b,* obtained for each participant, could be used to describe individual performance and to investigate the possible effects of independent variables. The estimate of the trial block number on which precurrent response duration per correct would be equal to zero, even if this does not happen during the session, could be obtained by dividing *b* by a *(bla).*

In order to investigate possible relationship's between individual response patterns and Equation 1 parameters, Pearson correlation coefficients among *b,* a, and *bla* were calculated across the parameters obtained for ali 58 participants. The correlation coefficient between a and *b* was equal to .96 *(p* =.000), indicating that participants whose precurrent response duration per correct was higher on the first trials (high *b)* showed higher rates of decrease in precurrent response from trial to trial (high *a).* The correlation coefficient between a and *bla* was equal to -.33 *(p* = .01), indicating that participants whose rate of decrease in precurrent response from trial to trial was higher (high a) needed a smaller number of trials to stop looking up the auxiliary screen (Iow *bla).* The correlation coefficient between *b* and *bla* was -.17 *(p* = .21), indicating that there was no linear relationship between precurrent response duration on the first trials *(b)* and the number of trials to stop looking up the auxiliary screen *(bla).*

These correlations among Equation 1 parameters suggest that an adequate global measure of performance would have to take into account the possible interactions among them. Such global measure should allow direct comparisons between, for example, response patterns showing high a and *b* with low *bla* and others showing low a and *b* with high *bla.* The area of the function derived from Equation 1, which is equal to *f:i212a,* would be a good candidate for a global measure of performance, for it would give the estimated sum of precurrent response duration per correct (I.e., for each correct current response) up to the trial on which precurrent response duration would be zero. This measure would make possible direct comparisons of different response patterns, for it could be loosely interpreted as the estimated duration of training necessary for correct current responding without precurrent responses (Le., for memorizing or automatizing).

Effects ot task camplexity. Equation 1 parameters were calculated with the data from each session of each participant in Conditions 2 to 5. Determination coefficients *(F)* ranged from .36 (P3 in Condition 4) to .91 (P6 in Condition 5) with a mean of .65 *(SO=* .13, *N* = 96), suggesting a good fit of the equation.

According to the design adopted, comparisons should be made between performance observed in Conditions 2 versus 3, which differed with respect to the average programmed frequency of reinforcement, in Conditions 2 versus 4, which differed with respect to the programmed probability of reinforcement given a shape, and in Conditions 4 versus 5, which differed with respect to the programmed probability of reinforcement given a position. Figure 5 shows the regression line (Equation 1) of duration of precurrent responses per correct as a function of the logarithm of trial blocks, for each session of each participant, obtained in Conditions 2 to 5. Each graph includes the regression lines, which show the areas of the functions, calculated for each participant in one group. The area of the function was larger in Condition 3 (M = 2.09, SD = 0.97), when compared to Condition 2 (M = 1.25, SD = 0.63), for 20 participants, indicating a systematic effect of the average programmed frequency of reinforcement. The area of the function in Condition 4 (M = 4.21, SD = 2.34) was larger than in Condition 2 for all participants, showing a systematic effect of programmed probability of reinforcement given a shape, and larger than in Condition 5 (M = 2.88, SD = 1.49) for 15 participants, suggesting some effect of programmed probability of reinforcement given a position. It can also be observed in the figure that, whereas in Groups 1 and 4, for which Character-Set 4 was used in Condition 4, the area of the function was larger in Condition 4 than in Condition 5 for 11 out of 12 participants, this occurred only for 4 out of 12 participants in Groups 2 and 3, for which Character-Set 5 was used in Condition 4. These results suggest that interaction effects between character sets and conditions may have occurred.

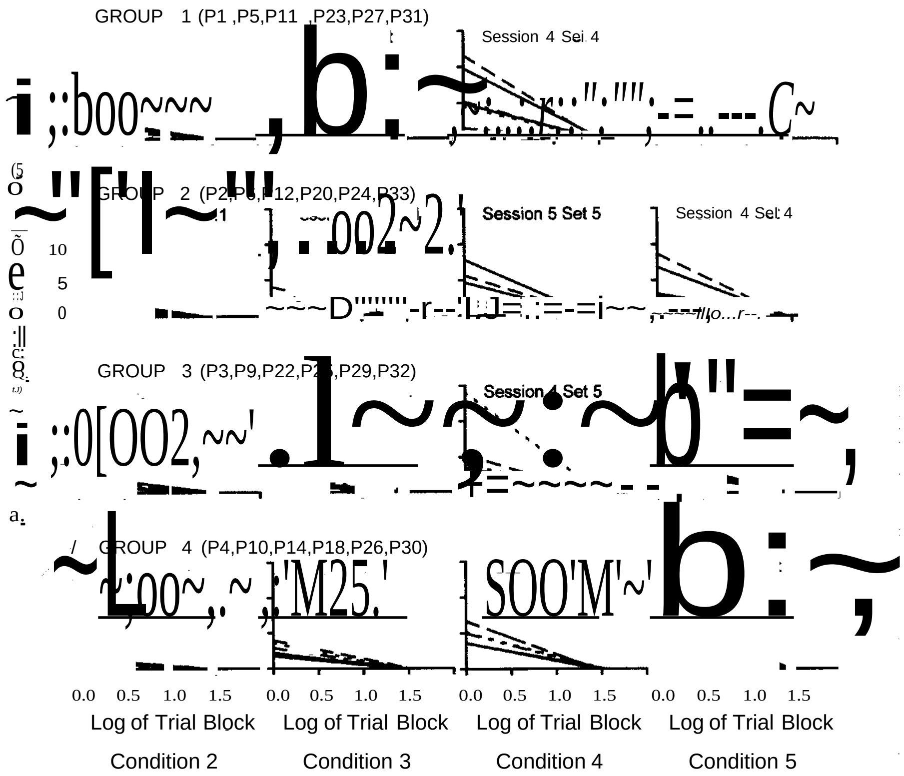

Figure 5. Regression lines, obtained from Equation 1, calculated for each participant in Conditions 2 to 5 in Experiment 3. Each graph shows regression lines for participants in one group, indicating the session number and character set used.

A detailed analysis of the character sets used (see Figure 3) indicated that one character (L) was used in Character-Sets 1 and 2 whereas another one (~) was used in Character-Sets 1, 3, and 5. These repetitions of some characters in different sets were not intended and may be related to the observed differences between Groups 1 and 4, on one hand, and Groups 2 and 3, on the other, concerning their performance in Conditions 4 and 5. The second repetition of one character in Character-Set 5 may have made it easier than Character-Set 4, which had no character repetition, facilitating thus performance of Groups 2 and 3 in Condition 4, when compared to their performance in Condition 5 using Character-Set 4. The repetition of one character in both Character-Sets 2 and 3 may have canceled out possible effects of sets across Conditions 2 and 3.

As an attempt to clarify these results, a three-way ANOVA, comparing the values of area as a function of Condition (2 through 5), as a within-subject factor, Character Set (sequence 2, 3, 4, 5 or 3, 2, 5, 4) and Session (sequence 2, 3, 4, 5 or 3, 2, 5, 4), as between-subject factors, was calculated. The ANOVA indicated, with an alpha level of .05, a significant effect of condition, $F(3, 60) = 32.39$, $p = .00$, and nonsignificant effects of character set, $F(1, 20) = .23$, $p = .64$, session, $F(1, 20) = .91$, $p = .35$, and interactions $ps > .73$ for all of them). A Tukey's test for multiple comparisons yielded a minimum difference of 0.86, with an alpha level of .05, and indicated significant differences between Conditions 2 (M = 1.25) and 4 (M = 4.21), and Conditions 4 and 5 (M = 2.88). The difference between Conditions 2 and 3, which was equal to .84, was very close to the minimum significant difference. These results suggest systematic effects of the variables related to task complexity and no systematic effects of session or character set, corroborating the proposal of analyzing task complexity based on the quantification of programmed contingencies of reinforcement.

With the purpose of testing whether individual differences in Condition 1 (Session 1) were related to performance in the other conditions, the Pearson correlation coefficient was calculated between the values of area of the function obtained in Condition 1 and the mean of values of area obtained in Conditions 2 to 5. This correlation coefficient was equal to .62 ($p = .00$), indicating that those participants who had higher performance in Condition 1 also showed higher performance in the other conditions.

General Discussion

Effect of Training on Precurrent Behavior

The results from Experiments 1, 2, and 3, showing decreases in the duration of this type of precurrent behavior, which in some cases stops occurring, as a function of increasing training, serve as experimental demonstrations of the conceptual and theoretical analyses described earlier. The observed regularity according to which precurrent behavior decreased, in all three experiments, despite differences in the tasks used, may encourage the adoption of this type of procedure in the investigation of the variables that influence precurrent responses, as done in Experiment 3. These results suggest that a negatively accelerated function (e.g., Equation 1) may describe well the relationship between duration of this type of precurrent behavior and training, and that individual differences in performance are partially maintained across conditions.

It should be noted, however, that this negatively accelerated shape, which is very similar to the classic learning curve, first described by Thorndike (1898), may be a consequence of the type of task and the measures used in the experiments (cf. Skinner, 1953, p. 53). If a less complex task were used, for example, only one pair of shape-number, ar the function were plotted for each pair separately rather than for the average of each trial block, or the frequency of precurrent responses were adopted as the main measure, individual data could have the shape of a step function.

In the case of the experiments reported here, in which no restriction was imposed upon the time participants could look up the auxiliary screen, duration of precurrent behavior per correct seemed to be a more interesting measure than frequency, for data based on frequency would differ greatly, across trials and participants, with respect to duration. Had the duration of the looking-up-the-auxiliary-screen response been kept constant, a frequency analysis would have been completely adequate. A frequency analysis would also be more relevant if the numbers to be memorized were larger (e.g., 20-digit numbers), in which case participants would probably look up the auxiliary screen several times for each pair on a given trial block. This would differ from the results obtained in the present experiments where, in general, participants looked up the auxiliary screen only once for each pair on each trial block.

It may be relevant to call attention to the fact that, in the present procedure, the frequency and duration of auxiliary stimuli presentation were completely under the control of the participants, which is not necessarily the case in precurrent contingencies. On the contrary, it seems that in the most commonly adopted teaching and experimental contingencies, the presentation of auxiliary stimulus, including fading procedures, is controlled by the instructor or experimenter, not by the learner. As mentioned previously, in some experiments, the prohibition of precurrent responses by the experimenter, after phases in which these responses were required, decreased the frequency of precurrent responses but was accompanied by decreases in correct current responding (e.g., Parsons et al., 1981; Torgrud & Holborn, 1989). These results differed from the ones reported here (and those from Polson et al., 1997), which showed correct current responding despite decreases in precurrent responses. Considering that the experiments cited above arranged precurrent contingencies that did not prevent transfer of stimulus function (as it is the case in typical observing behavior procedures), these differences in result could be partially explained by the fact that, in the procedures adopted here (and in Polson et al., 1997), participants, rather than the experimenter, could control when and whether they would stop emitting the precurrent response. This level of controlability of auxiliary stimulus presentation, as one of the characteristics of teaching procedures, may have a substantial effect upon performance and should be systematically investigated.

The present results would also call attention, in the context of an operant theory, to the possibility of decrease in frequency (and/or duration) of responses correlated with increased probability of reinforcement. Precurrent responses decreased (and even stopped occurring) despite the fact that they increased the probability of reinforcement for current responses. This apparent contradiction could be explained by considering that the probability of emitting the correct current response in the absence of auxiliary stimuli increases as the training proceeds (which is the mirror function of the one presented here, i.e., the increase in control by the stimulus that occasions current responding). Therefore, the probability of reinforcement for the current response is no longer higher in the presence, when compared to the absence, of auxiliary stimuli. In the present experiments the instructions explicitly asked the participants to reduce responses to the auxiliary screen, by emphasizing that their task was to memorize the material and by arranging explicit contingencies, through point delivery, for these responses to stop occurring. One could speculate, however, that neither explicit instructions nor specific contingencies (positive or negative) is a necessary condition for decreases in precurrent behavior. Other variables that may influence the decrease of this type of precurrent behavior should be analyzed, such as reductions in reinforcement delay and response cost, as has been suggested by Skinner (e.g., 1957, p. 436), with the expression *labor-saving,* when theorizing about some of the conditions that would transform public into "covert" behavior.

The investigation of this type of precurrent behavior may also raise questions relevant to typical research concerned with the effect of practice upon the time to execute different tasks. Recent research (e.g., Newell, 1991; VanLehn, 1996) has suggested that the decrease in time to perform a task with increasing practice, the learning curve, is best described as a power function. As mentioned previously, the decrease in precurrent behavior cannot, at least without theoretical adaptations, be adequately described as a power function because this type of behavior may stop occurring. Theoretically, the function would have to allow for values equal to zero, which a typical power function does not. This would imply that the learning curve, as typically conceived, would have to be decomposed in more than one function, which would describe at least two different kinds of responding: the decrease of precurrent behavior and the decrease of current (or other types of precurrent) behavior.

Task Complexity and Precurrent Behavior

Results from Experiment 3 showed that increases in the number of different responses to be learned in the total set, per pair, and per position, were associated to decreases in performance, corroborating those found in the literature concerned with task complexity, despite their theoretical and procedural differences (e.g., Bãuml, 1997; Brehmer, 1992; Coren & Ward, 19~9; Ebbinghaus, 1885/1964; Kerstholt, 1992; Payne, 1982; Postman, 1972; Ratcliff, Clark, & Shiffrin, 1990; Simon, 1972, 1974; Sündstrom, 1987). It may be relevant to notice that the measure of performance adopted in the present experiment, that is, the area of the function, which can be interpreted as the total duration of precurrent response necessary to learn each current response, is similar to those used in the first experimental investigations of memory (cf. Ebbinghaus, 1885/1964), which measured the time of study (or number of repetitions) necessary to memorize each item of a list.

The obtained results are compatible with an analysis of task complexity based upon the quantification of programmed contingencies of reinforcement. According to this analysis, changes in the number of different responses to be learned in the total set, per pair, and per position, may be interpreted, respectively, as changes in the average programmed frequency of reinforcement in the task and the programmed probabilities of reinforcement for a response given each discriminative dimension (i.e., shape and position). This type of analysis may be, in principle, applicable to other tasks. In reading (or writing), for example, the programmed probability of reinforcement for the emission of certain sounds (or writing certain letters) in the presence of certain letters (or sounds) may vary considerably in the same language and in different languages. This quantification of the programmed contingencies may provide a more refined analysis of what some authors have called irregular letter-phoneme relations (e.g., Alessi, 1987).

Although the analysis presented here may serve to demonstrate the possibility of describing some aspects of task complexity as the quantification of programmed contingencies of reinforcement, some theoretical issues should be discussed. First, it should be pointed out that despite the fact that the programmed consequence for correct/incorrect responses' occurred, in Experiment 3, after the participant had typed five characters, in the analysis presented here each character to be typed was interpreted as a different response, considering that any change in any of them could change the consequence, that is, each character to be typed was a functional unit according to the programmed contingencies. It should be noted that this interpretation does not assert anything about the actual functional units that may have been formed throughout the experiment (for a discussion, see Catania, 1992, p. 124).

Secondly, although this kind of analysis may be applicable to several different types of tasks, such as serial learning tasks and simultaneous discrimination procedures, it is clearly not applicable to some types of task. This is particularly the case of tasks in which the complexity changes by altering the type of movement to be made. In the task used in the present experiments, the type of movement to be made (pressing keyboard keys) was somewhat trivial, in the sense that all participants could easily make them before the experiments, and was not varied across conditions. Therefore, the complexity of the task was not related to how to make certain movements but rather to what movements should be made (i.e., what keys should be pressed) under what situations (i.e., in the presence of what shapes and positions). This type of complexity might be named *discriminative complexity* in order to distinguish it from something like *topographical complexity*, this latter being related to how to make the correct movements. The task used would then be characterized as having very low topographical complexity and varied discriminative complexity. In ballet dancing, or tennis playing, one could find examples of situations with low discriminative complexity and high topographical complexity, that is, people may know what movements they should make in what situation but they cannot make the correct movements or make them with the appropriate speed or accuracy (it may be interesting to notice that it does not seem possible to describe topographical complexity independently of the repertoire of the individual). Many tasks would fall between these extremes, that is, they could be described as having some degree of both types of complexity. The procedures to remediate performance problems would differ according to the type of complexity, for whereas performance problems related to topographical complexity would require, for example, shaping techniques, discriminative complexity would involve, probably, fading procedures.

References

ALESSI, G. (1987). Generative strategies and teaching for generalization. *The Analysis of Verbal Behavior,* 5,15-27.

BAuML, K. H. (1997). The list-strength effect: Strength-dependent competition or suppression? *Psychonomic Bulletin & Review,* 4, 260-264.

BICKEL, W. K., HIGGINS, S. T., & HUGHES, J. R. (1991). The effects of diazepam and triazolam on repeated acquisition and performance of response sequences with an observing response. *Journal of the Experimental Analysis of Behavior,* 56, 217-237.

BLOUGH, D. S. (1959). Delayed matching in the pigeon. *Journal of the Experimental Analysis of Behavior,* 2,151-160.

BREHMER, B. (1992). Dynamic decision making: Human control of complex systems. *Acta Psychologica,* 81, 211-241.

BRUNER, A., & REVUSKY, S. H. (1961). Collateral behavior in humans. *Journal of the Experimental Analysis ot Behavior,* 4, 349-350.

CATANIA, A. C. (1992). *Learning* (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.

COREN, S. E., & WARD, L. M. (1989). *Perception and information.* New York: Academic Press.

D'AMATO, M. R., ETKIN, M., & FAZZARO, J. (1968). Cue-producing behavior in the capuchin monkey during reversal, extinction, acquisition, and overtraining. *Journal of the Experimental Analysis of Behavior,* 11, 425-433.

DEESE, J. (1958). *The psychology of learning.* New York: McGraw-Hill.

DINSMOOR, J. A. (1983). Observing and conditioned reinforcement. *Behavioral and Brain Sciences,* 6, 693-728.

EBBINGHAUS, H. (1964). *Memory: A contribution to experimental psychology.* New York: Dover. (Original work published 1885).

ECKERMAN, D. A. (1970). Generalization and response mediation of a conditional discrimination. *Journal of the Experimental Analysis of Behavior,* 13, 301-316.

HACKMAN, J. R. (1969). Toward understanding the role of tasks in behavioral research. *Acta Psychologica,* 31, 97-128.

HEARST, E. (1962). Delayed alternation in the pigeon. *Journal of the Experimental Analysis ot Behavior,* 5, 225-228.

HODOS, W., ROSS, G. S., & BRADY, J. V. (1962). Complex response patterns during temporally spaced responding. *Journal of the Experimental Analysis of Behavior,* 5, 473-479.

JANS, J: E.,.~ C~TANIA, A. C. (1980). Short-term remembering of discriminative stimuli In pigeons. *Journal of the Experimental Analysis of Behavior* 34 177-183.

KERSTHOLT, J. H. (1992). Information search and choice accuracy as a function ~task complexity and task structure. *Acta Psychologica,* 80, 185-197.

LATIE~, V:. G:, ~EISS, .B., C~ARK, R. L., & REYNOLDS, M. D. (1965). Overt mediating behavior during temporally spaced responding. *Journal of the Experimental Analysis of Behavior,* 8, 107-116.

LATIES, V. G., WEISS, B., & WEISS, A. B. (1969). Further observations of overt "mediating" behavior and the discrimination of time. *Journal of the Experimental Analysis of Behavior,* 12,43-57.

MUELLER, K. L., .& ~INSM~OR, J. A. (1984). Testing the reinforcing properties of S-: A replication of Lieberman's procedure. *Journal of the Experimental Analysis of Behavior,* 41, 17-25.

NEWELL, K. M. (1991). Motor skill acquisition. *Annual Review of Psychology* 42,213-237.

OHTA, A. (1987). Observed responses maintained by conditional discriminative stimuli. *Journal of the Experimental Analysis of Behavior,* 48, 355-366.

OLIVEIRA-CASTRO, J. M. (1992). "Fazer na cabeça": Uso metafórico e negativo ["Doing in the head": Metaphorical and negative use]. *Psicologia: Teoria* e *Pesquisa,* 8, 267-272.

OLIVEIRA-CASTRO, J. M. (1993). "Fazer na cabeça": Análise conceitual, demonstrações empíricas e considerações teóricas ["Doing in the head": Conceptual analysis, empirical demonstrations and theoretical considerations]. *Psicologia USP,* 4, 171-202.

PARSONS, J. A., TAYLOR, D. C., & JOYCE, T. M. (1981). Precurrent self-prompting operants in children: "Remembering". *Journal of the Experimental Analysis of Behavior,* 36, 253-266.

PAYNE, J. W. (1982). Contingent decision making. *Psychological Bulletin, 92,* 382-402.

POLSON, D. A., GRABAVAC, D. M., & PARSONS, J. A. (1997). Intraverbal stimulus-response reversibility: Fluency, familiarity effects, and implications to stimulus equivalence. *The Analysis of Verbal Behavior,* 14, 19-40.

POLSON, D. A., & PARSONS, J. A. (1994). Precurrent contingencies: Behavior reinforced by altering reinforcement probability for other behavior. *Journal of the Experimental Analysis of Behavior,* 61,427-439.

POSTMAN, L. (1972). Transfer, interference and forgetting. In J. W. Kling & L. A. Riggs (Eds.), *Woodworth & Schlosberg's Experimental Psychology* (pp. 1119-1132). New York: Holt, Rinehart and Winston.

RATCLIFF, R., CLARK, S., & SHIFFRIN, R. M. (1990). The list-strength effect I: Data and discussion. *Journal of Experimental Psychology: Learning, Memory and Cognition,* 16, 163-178.

RYLE, G. (1949). *The concept of mind.* London: Hutchinson,

SHIMP, C. P., & MOFFITT, M. (1977). Short-term memory in the pigeon: Delayed-paired-comparison procedures and some results. *Journal of the Experimental Analysis of Behavior,* 28, 13-25.

SIMON, H. A. (1972). Complexity and the representation of patterned sequences of symbols. *Psychological Review,* 79, 369-382.

SIMON, H. A. (1974). How big is a chunk? *Science,* 183,482-488.

SKINNER, B. F. (1953). *Science and human behavior.* New York: The Macmillan Company.

SKINNER, B. F. (1957). *Verbal behavior.* Englewood Cliffs, NJ: Prentice-Hall.

SKINNER, B. F. (1968). *The technology of teaching.* Englewood Cliffs, NJ: Prentice-Hall.

SKINNER, B. F. (1969). *Contingencies of reinforcement: A theoretical analysis.* Englewood Cliffs, NJ: Prentice-Hall.

SÜNDSTROM, G. A. (1987). Information search and decision making: The effects of information display. *Acta Psychologica,* 65, 165-179.

THORNDIKE, E. L. (1898). Animal intelligence: An experimental study of the associative processes in animais. *Psychological Review Monograph Supplements,* 2(No. 4), 109.

TORGRUD, L. J., & HOLBORN, S. W. (1989). Effectiveness and persistence of precurrent mediating behavior in delayed matching to sample and oddity matching with children. *Journal of the Experimental Analysis of Behavior,* 52,181-191.

VANLEHN, K. (1996). Cognitive skill acquisition. *Annual Review of Psychology,* 47,513-539.

WOOD, R. E. (1986). Task complexity: Definition of the construct. *Organizational Behavior and Human Decision Process,* 37, 60-82.