Measuring Response Rates in Associative Skill Development

Carl Binder

Unpublished Draft Doctoral Working Paper
Experimental Psychology Department – Harvard University
Cambridge, MA

1978

This unpublished document combines references to research and application seldom cited in the literature of Precision Teaching or Fluency-based Instruction. Scanned from an original working draft, it was never published or presented as is, but used in preparation of several published papers. Its perspective and combination of topics were shaped by the scholarly need to connect some of our early work in Precision Teaching with basic research in experimental psychology. Its purpose was to pull together strands of experimental learning research in which the time dimension –latency, rate of response (frequency), or response duration – was included in the measurement of learning and performance.

The first chapter covers a diversity of sources under a common term, “associative skills.” It highlights research and theory from multiple sub-fields that might be relevant to combinations of motor and verbal skills, including discriminations and combinations, similar to those taught in Precision Teaching classrooms. It summarizes relevant effects and findings.

The second chapter explores implications of using Skinner’s variable, *rate of response*, as a measure of skill *proficiency* – a term used by Precision Teachers at the time. It draws on precedents in learning research, and presents some pilot research from our lab classroom.

The third chapter traces the evolution from rate of response in the laboratory to standard celeration charting and Precision Teaching in the classroom. It introduces the term *fluency,* and summarizes our research program in the Behavior Prosthesis Lab within the framework we developed of “four kinds of ceilings” limiting student performance that emerged as we measured response rates (i.e., behavior frequencies) in the classroom. We hope that this backgrounder might be helpful to scholars and practitioners in Precision Teaching and related fields.

NOTE: Please pardon use of the term *retarded* as in *retarded students*, or *severely retarded*, in chapters II and III, written in an era when that term was accepted practice among educators and scholars, an advance from previously used terms, *idiot*, *imbecile*, and *moron*.

Contact the author at carlbinder@fluency.org

Chapter I: Precedents for Associative Skill Research

Experimental psychologists have traditionally distinguished between associative learning or verbal learning, and motor learning or skill acquisition. The categories of research and theory which these terms denote are not entirely exclusive of one another in experimental practice. Yet scholars and scientists writing about one field rarely discuss work conducted in the other. Most books on human learning deal with one or the other of these topics. When both areas are considered within a single volume (e.g., Marx, 1972), associative learning and motor learning are covered in separate sections, usually written by separate authors.

A few writers have tried to bridge the gap. For example, Fitts (1964) argued that:

> Sharp disctinctions between verbal and motor processes, or between cognitive and motor processes serve no useful purpose . . . since the processes which underlie skilled perceptual-motor performance are very similar to those which underlie language behavior as well as those which are involved in problem solving, and concept formation. (p. 243)

From a behavior-analytic standpoint, demonstration of any kind of discrimination learning requires the existence of a well-differentiated response. Such indicator responses, viewed in isolation, may appear too trivial to interest experimental researchers. However, viewed in the context of any but the simplest and most rarefied human performance, motor components assume a critical role which may either inhibit or facilitate the process of associative learning.

Underwood (1966) observed that:

> We rarely study the learning of the skill to execute a particular verbal response. Learning to give the proper pronunciation to words in a foreign language would be an illustration. Or suppose we required our subjects to respond with a particular frequency and intensity in paired-associate learning in addition to responding to the appropriate stimulus. Under these circumstances we would be more nearly studying skill learning. (p. 495)

Similarly, Lenneberg (1964) discussed "the relation of speech as a motor skill to language as a psychological skill."

Reading, writing, basic computational skills, speech, and the playing of musical instruments, all involve interactions between motor and associative learning. In the terms of operant conditioning, such skills require the behaver to emit sequences of topographically complex responses under the control of rapidly changing stimulus configurations. In terms of the traditional categories of study, they represent an intersection between motor and associative processes. The term <u>associative skill</u> seems an appropriate label for the category of human performance which includes these classes of behavior and a variety of other sub-categories as well.

With the possible exception of reading, experimental scientists rarely study any of these skills. By studying behavioral paradigms which represent processes assumed to underlie these skills, scientists hope to discover relevant functional relations. However, there are probably synergistic relationships among the components of associative skills in the form of mutual growth dependencies. Deficits in certain components may constrain learning of others, just as development of certain components may facilitate learning of others. Thus, study of the parts in isolation may be misleading in a variety of ways. For example, isolated processes may become objects of study in themselves in ways which divert research from issues of practical significance. Rote memory tasks, for instance, bear little resemblance to human behavior in vivo, yet they have assumed a dominant role in research and theory on human learning. Only by viewing such simple associative processes within the context of ongoing skilled performances such as reading, language-learning, or computation can scientists estimate their relative importance as components of more complex processes.

Moreover, study of simple processes in isolation may actually complicate our understanding of composite associative skills. Variables producing significant effects in the laboratory study of isolated processes (e.g., <u>meaningfulness</u>, in the study of paired-associate learning) may fade into relative insignificance in the study of composite associative skills. The effects of repeated practice, for example, might override the effects of such minor dimensions as meaningfulness in the performance of associative skills. Direct study of associative skills themselves might clarify our understanding of the relative importance of various independent variables, thereby leading to a more parsimonious account of the subject. Even apart from more theoretical concerns, the practical implications of such analyses seem to warrant basic research efforts.

Motor Learning and Associative Learning

Emphasizing the points at which the two areas of research intersect, the following is a selected review of some of the paradigms, measurements, and major areas of concern in the fields of motor learning and associative learning. For the most part, the review is neither critical nor interpretative. Rather, it represents a search for precedents and general findings which may be relevant to the analysis of associative skills.

Major Paradigms

Associative learning. Originally used by Ebbinghaus (1885), serial learning of nonsense syllables was the most frequently studied verbal learning task until the early 1950's (Hall, 1971). In serial learning experiments, subjects learn lists of items, and then are tested on their ability to recall the lists in the correct order. Experimenters vary specified characteristics of the list items and the length of the lists, assessing the effects of such manipulations on recall. Like free recall tasks, in which subjects are read lists of items which they then must recall in whatever order they can, list-learning tasks have little direct relationship to most forms of associative skill performance.

Paired-associate tasks, on the other hand, bear a closer resemblance to many associative skills. In general, these tasks involve the learning of stimulus-response pairs, the elements of which may vary across a wide range of dimensions. Stimulus elements may be visual, auditory, or even tactile; meaningful or nonsensical; verbal or nonverbal. Responses may be interchangeable with stimuli or not; verbal or motor; nonsensical or meaningful. In short, studies of paired-associate learning may include almost any form of associative pairing that an experimenter can invent.

For this reason, especially in recent years, the paradigm has become the one most commonly used in the study of human associative learning. Most experimenters present pairs of associates in a trial-by-trial fashion, and as a rule, measure learning in terms of percentage correct (Hall, 1971).

Motor learning. McGeoch (1927) wrote one of the first reviews of the motor learning field. In a more recent review Adams (1972) outlined four categories of motor learning tasks. First, simple gross- and fine-motor exercises allow researchers to examine the effects of such variables as fatigue, warm-up, and task requirements on rate and intensity of responding. Second, simple perceptual-motor tasks such as lever-positioning and line-drawing (specified length) allow experimenters to examine the effects of such variables as kinesthetic feedback and post-performance knowledge of results on acquisition and retention of accurate responding. Third, continuous tasks involving pursuit-rotors and other forms of motor-tracking are studied, largely because of their relationship to such skills as air-plane-piloting and certain machine-shop operations. Experimenters measure percentage of time on target in continuous motor tasks. Finally, perceptual-motor tasks involve discrete responses at the onset or termination of different kinds of signals. In these tasks, both reaction time and percentage of correct responses are measured. This last category of task is related in obvious ways to paired-associate learning. In fact, some such tasks have been referred to as "motor paired-associate" tasks (e.g., White, et al., 1960).

A thorough review of the literature and history of the motor learning area (Irion, 1966, 1969) reveals that paradigms and procedures have never been standardized, and that the field exhibits a remarkable lack of continuity. Fleishman (1966) illustrates a variety of the mechanical devices which experimenters have used in the study of motor learning, and it is clear that this is just a sampling. A major reason for such diversity and discontinuity may be that researchers have very often conducted studies with specific tasks and particular training concerns in mind (Wolfe, 1951; Irion, 1966). Consequently, they have designed apparatus and experiments to meet very specific requirements for tasks which may become obsolete within a matter of months or years. Much of the work in motor learning has arisen in the context of human engineering, and especially under the auspices of the military and other government agencies. As the tanks and radar change, so change the simulators and research questions. There have, nonetheless, been a number of common research topics and general findings which invite comparison with those of associative learning research.

Intersections and Common Findings

Transfer of training. Transfer of training has been a major focus in both associative and motor learning research. Postman (1971) and Jenkins (1963) provide comprehensive reviews of transfer of training in associative learning. Irion's (1966, 1969) historical accounts and Briggs's (1969) review are probably among the best reviews of the topic in motor learning. Much of the work in both fields is either explicitly or implicitly concerned with instructional technology, and certain workers (e.g., Gagne, 1970) deal almost exclusively with the implications of basic research for instructional design.

In the verbal learning field, a wide variety of both positive and negative transfer effects have been functionally related to a broad range of independent variables. The list of robust associative transfer effects is nearly endless. And the catalog of variables which produce these effects is also extensive. It includes dimensions of stimuli, responses, and procedures which are of practical significance, as well as some which are not. Mediated transfer effects are especially interesting. Horton and Kjeldergaard (1961) demonstrated eight different arrangements of three stimulus-response pairs which can produce significant effects. For example, subjects who learn to associate both A-B and A-C, will be able to associate B-C as an effect of mediated transfer. These paradigms, which have been extended to multiple stages of associative transfer, are intriguing to researchers because they seem to provide simple models for inductive and deductive reasoning. Perhaps more than any other topics of associative learning research, mediated transfer paradigms seem to confirm the theories of Aristotle and the British associationists about the nature of thinking (Herrnstein and Boring, 1965).

In motor learning, transfer of training seems nearly always to work in one direction (Irion, 1966). Practice on tasks which resemble the transfer task will frequently produce a positive transfer effect. However, except in the case of motor paired-associate tasks and other perceptual-motor tasks with strong associative components, negative transfer seldom occurs (Irion, 1966).

In two different multidimensional tracking tasks, where subjects were required to stay on target with respect to movement in two directions, Bilodeau and Bilodeau (1954) found that improving the performance of a less-proficient component, even at the expense of practice on a more-proficient component, could boost overall task proficiency.

In two different experiments, Gagne (Gagne and Foster, 1949; Gagne, Baker and Foster, 1950) investigated the effects of practice on components of a discrete motor task. The apparatus included a panel of four lights with a column of red lights on the left and a column of green lights on the right (or vice versa). Beneath the light panel, a square array of four switches was arranged so that a given row of switches corresponded to a given color of light. In one study, the researchers found that simple reaction-time training on the switching-response facilitated performance on the composite task. In the other study, training on either the color or position component of the light-switch association facilitated performance on the composite task. Data showed the color component to be more difficult to master than the position habit. And as in the Bilodeau and Bilodeau (1954) study cited previously, training on the more difficult component had a greater overall effect than training on the less difficult. Gagne et al. (1950) concluded that "preliminary training on a given discrimination reduced the final task errors associated with discriminations equally or less difficult, but does not reduce those errors associated with a more difficult discrimination." This conclusion seems particularly relevant to the study of associative skills where deficits in one component may inhibit development of others. For example, a deficit in the ability to distinguish among certain speech sounds may impair reading acquisition.

Citing an unpublished doctoral dissertation of R.E.Tulloss, Keller (1958) described a transfer of training sequence observed in the acquisition of Morse Code. Beginning telegraphers first learn to say a letter or digit in response to a visual dot-dash pattern. Next, they learn to visualize such a pattern upon hearing an auditory signal. They can then hear a signal and write the appropriate character by means of two mediating responses: visualizing the pattern and sub-vocally repeating the letter or number. With continued training the mediation may begin to occur by means of an overt (and subsequently covert) vocalization of the code in the form of "di-dah" sequences. Then the telegraphers become able to say or write the letter by means of a direct response to transmitted code. Subsequently, they learn to respond to larger units of code at a time, by saying or writing words or phrases. Those who learn to send as well as receive code must apply many of the same discriminative responses to the precise motor skill of using a telegraph key. The behaviors exhibited by practiced telegraphers provide excellent illustrations of associative skills.

Gagne (1970) has developed a conceptual framework and a research methodology for the study of transfer in what he calls <u>learning hierarchies</u>. A learning hierarchy is a sequence of skills or "knowledges" which lead to a terminal objective, and which are empirically determined to be both necessary and sufficient prerequisites for the emergence of the terminal skill. The learning hierarchy formulation provides a useful framework for the study of motor skills, associative learning, and in particular, associative skills.

<u>Retention</u>. Ebbinghaus (1885), the acknowledged father of the field of verbal learning, studied the retention of verbal associations. Early verbal learning researchers were functionalists who sought to derive learning and retention curves as a function of various experimental manipulations (Underwood, 1966). Recently, the field has shifted to a more theoretical approach, emphasizing hypothetical constructs and mathematical models. Atkinson and Shiffrin's (1968) model of memory and its "control processes" has been especially influential during the last decade. Experiments testing the implications of this model have generated vast quantities of data. Whether such hypothetical-deductive research is an advance or a regression in methodology is not the focus of this discussion, although this writer prefers an atheoretical approach to the analysis of behavior. In any event, nearly

a century of research has revealed a great number of both short-term and long-term memory phenomena as a function of many different variables. And it is safe to say that decrements over time in discriminative responding, whether through processes of memory decay, interference, or extinction, are commonly-observed phenomena.

In the area of motor learning, however, the research has been neither extensive nor systematic. Whereas short-term memory curves for verbal learning and some forms of motor learning appear similar, there is argument about whether the underlying processes are the same (Adams, 1972).

With respect to long-term retention, researchers have found from the very beginning (McGeoch, 1927) that continuous motor-skills such as pursuit-rotor and tracking tasks are retained over very long periods of time. Eysenck (1960), for example,using a procedure which allowed a brief warm-up period at the beginning of the retention test, found nearly complete retention, after one year on a continuous perceptual-motor task. Such findings lend support to the adage that one never forgets how to ride a bicycle. On the other hand, retention of discrete motor responses appears more like that of associative learning: discrete motor tasks are much less likely to be retained than continuous tasks (Adams, 1972). This finding is not surprising insofar as discrete tasks usually contain significant associative components.

Naylor and Briggs (1961) point out that the number of variables which might affect motor retention is quite large and that very little work has been done on any one. They suggest four classes of variables for research in the area: type of task, learning parameters, retention parameters, and recall parameters. The first category, type of task, is particularly relevant in relation to the study of associative skills. Both stimuli and response topographies in associative skills may vary along dimensions of complexity and precision. They may or may not resemble the components of other skills. The discriminative components may be intrinsic to the response, as in printing letters, or they may be extrinsic, as in reading Chinese characters or operating computers. These aspects and others would be likely to affect the degree to which associative skills are retained over periods of time in which they are not performed. If we can assume that most skill acquisition occurs in sequences which require previously learned skills as components of terminal behaviors, then continued performance of these component skills is likely to maintain them. Thus the study of retention may be of little interest for researchers concerned with the development of instructional technology. Like so many laboratory conditions, extended periods of non-performance are rare in vivo.

Reminiscence and fatigue effects. Reminiscence is the phenomenon, observed primarily in perceptual-motor skills, whereby subjects perform significantly better immediately after a rest period than immediately before. It is associated with the experience of fatigue, and is one of the most commonly reported effects in the motor learning literature. Clark Hull may owe a good deal of his popularity in the 40's and 50's to his theoretical explanation of reminiscence. Hull (1943, 1945) postulated an intervening variable called reactive inhibition which was supposed to build up as an effect of continued motor performance, and to decay during periods of non-performance. Thus, performance would be suppressed more and more by reactive inhibition over the course of a practice period, only to return to its true or uninhibited level after the decay of inhibition during a rest. The reminiscence effect is the apparently discontinuous increment in performance observed after the rest period.

Reminiscence occurs in both continuous and discrete motor skills. Wright and Willis (1969), for example, found a greater reminiscence effect on the pursuit rotor in normal subjects than in "mental defectives." Otto and Fredericks (1963) observed reminiscence when they required subjects to practice printing numerals upside-down. This effect may be relevant to the study of associative skills, especially with respect to instructional design. The study of reminiscence and related effects may be helpful in the determination of optimum practice and assessment periods for associative skills. Research may also reveal relationships between these effects and levels of skill proficiency.

Reminiscence is related in an obvious way to the effects of massed versus distributed practice. And it may be a point at which associative learning and motor learning intersect in the performance of associative skills. Verbal learning studies have revealed only minimal effects of massed versus distributed practice on paired associate learning (Hall, 1971). Yet because associative skills contain both motor and associative components, reminiscence and fatigue effects are likely to be important factors in the study of such skills, and in the development of instructional procedures.

<u>Feedback past acquisition criteria.</u> Most researchers agree on the importance of contingent consequences in the acquisition of behavior. Operant conditioners discuss <u>reinforcement</u>, <u>punishment</u> and <u>extinction</u> in terms of observed functional relations. Many researchers in the field of motor learning refer to the <u>feedback</u> function. And most of what we know from studies of animal learning would lead us to think that contingent consequences are necessary for the maintenance of behavior beyond acquisition criteria. Yet a host of studies in both associative learning

(Butler and Peterson, 1965; Eimas and Zeaman, 1963: Goss, 1965: Richardson and Gropper, 1964) and perceptual-motor learning (Leonard and Conrad, 1963) support the opposite view.

In paired-associate tasks, both accuracy and reaction-time improve when subjects continue to pracrice without the benefit of feedback after they have met acquisition criteria. Leonard and Conrad (1963) taught their subjects a task resembling typewriting and found that the subjects more than doubled their rates of performance as a result of extra practice sessions during which they received no feedback.

Adams (1972), among others, has suggested that subjects provide their own feedback, or self-reinforcement, during post-acquisition practice. Whatever the formulation, it is clear that this common result in both the associative and motor learning fields has significant implications with respect to the development of associative skill proficiency beyond mere acquisition.

Studies of Reading

Clearly an associative skill, reading involves all the motor complexities of articulation (at least in its oral form), as well as a series of very fine eye movements. In addition, reading requires a great deal of associative learning. Probably because of its cultural significance, reading is an associative skill which experimental scientists have studied in some detail, for many years (Williams, 1973). Studies of reading may constitute the only large, coherent body of scientific research on an associative skill. Thus, a review of selected articles in the reading literature might help clarify the definition of associative skill.

Associative Components

One could argue that nearly all work in the field of paired-associate verbal learning is relevant to the study of reading. However, reading is a freely emitted, serial associative skill. Readers pace themselves in continuous performance. Thus, it may not be possible to draw any unconditional implications for the analysis of reading from the experimenter-paced trials procedures used in verbal learning research. Because of this limitation in verbal learning research, the most relevant studies concerning the associative components of reading are not experimental. Rather, they involve assessments across different groups and ages of subjects of such skills as naming pictures, naming colors, reading letters, syllables and words. Data collected in this way permit an analysis of the interactions between characteristics of the stimuli and of the responses which they control, in terms of differences in reading rates.

Doehring (1976), for example, assessed 150 school children across ten grades with timed tests in which subjects were required to read a variety of materials, including numerals, letters, syllables and words. Relative rates of reading these materials, and changes in rates across grade-levels, provided information about the relative degrees of control exerted by different kinds of stimuli. Such data allow the researcher to relate the associative components of reading, as it becomes more proficient, to lexical, syntactic, and other dimensions of verbal behavior. Doehring's work is important because it examines associative effects within the context of the associative skill performance. It should be possible to design analogous studies, including experiments, which permit the analysis of discriminative stimulus control within the context of other complex associative skills.

Motor Components and Interactions

Researchers have devoted a large amount of study to the motor components of reading, especially to eye-movements. Although studies of phonology and articulation are other potential sources of data on motor components of reading, the following discussion is limited primarily to the topic of ocular-motor research.

A number of researchers have studied the possibility that widening of the eye-span might improve the performance of reading and related perceptual-motor skills. Cattell (1886) conducted an experiment in which subjects read letters as they passed behind a slit. Subjects were permitted to see one or more than one letter at a time -- a condition called preview. He found that ability to see more than one letter at a time increased serial reading speed up to a preview limit of four or five letters, but that the major advantage was gained as soon as the second letter was visible. In a related study, Wagoner and Fitts (Fitts, 1964) required subjects to push one of five keys when light-points, moving downward in five different columns, passed behind a horizontal line spanning the five columns. Most of the advantage gained from allowing subjects to see more than one column at a time was achieved by preview of the next light in sequence. A preview of 12 oncoming lights was only slightly better than preview of one. Finally, in a review of eye-movement photography research, Taylor (1965) reported that training simply does not widen eye-span in a way that would affect reading speed. Physiological limits apparently constrain eye-span to about 2.7 words during normal reading. Speed-reading involves selective scanning and skimming rather than improved eye-span.

Thalberg (1967) examined the relationship between reading speed and retention. A group of slow readers and a group of fast readers read a 1500-word passage and took retention tests immediately afterward, as well as 24 hours later. Although the slower readers' retention scores were higher than those of the faster readers immediately after reading the material, the two groups showed equal retention after 24 hours. The obvious conclusion is that, all things being equal, it is better to read faster. These, and other findings, highlight the importance of investigating the factors responsible for reading speed. Speed of eye-movements is one such factor.

Speed of eye-movements (fixations and regressions) as well as eye-voice span have been topics of much research. There are at least two methods for directly measuring eye-movements. Electro-oculography, (Shackel, 1960) uses externally attached electrodes to measure the elec--trical potentials generated in the eye muscles. Ocular photography uses photographic apparatus to measure light, reflected from the eyes during movement.

Studies of eye-movement efficiency have shown that better readers make fewer fixations and regressions per word than poor readers. Schmidt (1966) found that college students developed greater efficiency, as measured by these two indices, as a result of a reading improvement course. Taylor, Frackenpohl, and Pattee (1960) found that first grade children made as many as two fixations per word, whereas 12th graders made one fixation every two words. Other researchers (Morton, 1964; Abrams and Zuber, 1972) have conducted related studies with similar results. The obvious implication is that it might be possible to increase reading speed by explicit eye-movement training.

Hoping to improve ocular-motor efficiency,reading researchers have built a variety of devices to measure and control the pace of reading (e.g., Goldiamond, 1962). However, in a study comparing the effectiveness of a tachistoscope, various controlled pacing devices, and simple "motivated" reading, Braam & Berger (1968) found that practice in reading paperback books with time-pressure motivation was as effective as any other method. Tinker (1967) also tested a variety of mechanical devices, and found them to be no more effective at improving rate of relevant eye-movements than motivated reading alone. Stennett, Smythe, Pinkney and Fairbairn (1972), using ocular photography, assessed the effects of practice on speed of eye-movements. They concluded that the simple skill of making relevant eye-movements is well-learned by the first grade, and that reading per se makes it efficient. Thus, they argue that efficient eye-movement is a result, not a cause of good reading.

Research on eye-movement efficiency illustrates the importance of investigating the associative and motor components of associative skills in the context of those skills.

The acquisition and development of associative skills may often involve interactions among their components. In the case of reading, the motor components apparently become more efficient as they are emitted under the control of textual stimuli. This is an important relationship, only apparent when the components are investigated in the context of the composite skill. Researchers interested in human associative learning might enhance their investigations by examining associative processes in the context of skilled performance.

Associative Skills and Covert Behavior

Discussions concerning human development, no matter what their theoretical standpoint, generally agree that the more abstract or covert human activities known as thinking, reasoning, and so on, develop as extensions of more concrete and overt behaviors. Although such covert activities may not be associative skills in the sense of having significant motor components, an understanding of associative skills should contribute to the analysis of covert behavior. Skinner's (1957) analysis of thinking in terms of the contingencies of overt verbal behavior may be the best illustration of such a contribution. Methods appropriate to the study of elementary written computation skills, for example, may also apply to an investigation of mental arithmetic. Thus, a definition of associative skill should apply, at least by analogy, to strictly covert behavior. And in the context of instructional design, it is clearly necessary to consider both associative skills and their covert extensions.

Summary and Conclusions

Associative skills are those which combine associative and motor components in a synergistic fashion. Examples include reading, writing, computational skills and musical performance. Experimental behavioral scientists have devoted very little research to such skills. Rather, researchers have usually studied motor and associative processes in isolation, or in extremely artificial forms. Intersections between currently existing bodies of associative and motor learning research suggest some possible approaches to the study of associative skills. Reading research is the major available illustration of the scientific study of associative skill.

Chapter II: Response Probability and Skill Proficiency

Introduction

In science, procedures and measurement operations delimit the extent of available information. The decision about what to measure and how to measure it defines, in part, the outcome of an investigation. Particularly if it occurs without an awareness of important alternatives, paradigm selection may severely restrict the advancement of scientific understanding.

One criterion for judging the goodness of an experimental paradigm is the degree to which it models a corresponding natural phenomenon. Insofar as a paradigm does reflect significant aspects of naturally occurring events, the conclusions which it yields can be extended to account for those events. With regard to measurement, the criterion is whether or not a given procedure measures the important dimensions of events as they naturally occur.

Typically, in the laboratory study of human learning, researchers measure percentage of correct responding or response latency in discrete trials procedures. There is a significant disparity between the behavior required in such procedures and human behavior as it occurs in ordinary daily activity. In particular, individuals participating in typical experiments are prevented from performing continuously, in response to complex sequences of environmental events. And they seldom have the opportunity to progress beyond the acquisition phase of learning. As a result, much of the experimental work on human learning has little relevance beyond the confines of the laboratory.

This chapter outlines the limitations of traditional approaches, and suggests an alternative approach to the measurement of human learning. It also reviews published data and presents some new data which illustrate the sensitivity of response rate as a measure of complex associative skills.

Limitations of Traditional Measurement Procedures

Neglect of the Time Dimension

A measurement-imposed ceiling. In both associative learning research and common educational practice, the usual measure of learning is percentage correct. Researchers and educators alike emphasize accurate performance, often as the sole criterion of learning. Many researchers seem to assume, without question, that percentage correct is the appropriate measure of learning. The term overlearning is indicative of this implicit assumption. It suggests a ceiling, beyond which it is impossible to measure learning. And this ceiling is 100% correct. So-called overlearning trials allow for learning beyond the 100% acquisition criterion. As Hall (1971) has pointed out, "It is recognized that the criterion of learning is arbitrary; consequently so-called over-learning trials merely provide the subject with the opportunity to achieve a more stringent criterion" (p. 429). But Hall's use of the term "criterion" is inappropriate in this context since there is no way of specifying a standard beyond 100% by which to make a criterion judgement. Thus, research on overlearning typically provides only indirect measures of the effects of added trials by means of retention or transfer scores. And the overlearning itself is operationally defined in terms of a specified number of trials.

It is clear that the amount learned per trial differs for fast and slow learners. Kruger (1929) attempted to quantify overlearning within subjects by setting the number of trials required to achieve 100% accuracy as a unit, and then defining the value of additional trials proportionally. He found, however, that a given number of overlearning trials did not produce the expected proportional increase in retention. It therefore seems impossible to scale overlearning in terms of percentage correct.

Latency: measuring beyond accuracy criteria. Osgood (1946) was among the first to design verbal learning experiments "to permit a latency measurement as a more sensitive indicator of habit strength than the mere presence or absence of the correct response" (p. 283). Although most researchers have ignored his early suggestion, experimenters who have measured latencies in associative learning research have nearly always arrived at the same conclusions.

Peterson (1965) found, in a paired-associate task in which subjects responded with digits to nonsense stimuli, that latencies decreased after the last error on an individual pair with no reliable trend in latency prior to the last error. Other studies have produced similar results (Keller, Thomsom, & Tweedy, 1967; Milward, 1964; Suppes, Groen & Schlag-Ray, 1966; Judd & Glaser, 1969; Hall & Wenderoth, 1972).

Judd & Glaser (1967) reported that "latency did not covary with increasing response probability during acquisition, but during overlearning latency showed a large and consistent decline" (p. 30). The authors' use of the term "response probability" in this context deserves note in that it is equated with percentage correct. It seems clear that with the addition of the time dimension as a measurement scale, the concept of response probability should expand to include immediacy of response, as well as simple accuracy. I will deal with this suggestion in more detail in subsequent portions of this chapter.

Judd & Glaser (1969) observed that latencies prior to the last error on a particular paired-associate were independent of the subject's learning rate, measured by trials to criterion. On the other hand, latencies after the last error on a given pair were an increasing function of learning rate. That is, subjects who met the accuracy criterion more quickly also decreased their latencies after the last error more quickly. This finding, like that of Kruger (1929) cited previously, poses serious questions for experimenters who define overlearning in terms of a fixed number of extra trials beyond 100% accuracy. They may be assuming that a given number of trials produces a given amount of overlearning. Yet this is clearly a false assumption, associated with the limitations of the percentage correct scale. Neglect of the time dimension in measurement procedures precludes the possibility of anything but operational definition and indirect measurement of learning beyond 100% accuracy.

A series of perceptual-motor skill experiments led Fitts (1964) to conclude that "skill learning, as measured by reaction time, continues far beyond the point where most verbal learning studies are terminated" (p. 265). And Theios (1973), discussing research on memory and cognitive processes, outlined the principles for using reaction time as an "indicator" of cognitive processing, memory scanning, attentional mechanisms, and similar hypothetical constructs. In fact, experimenters in the fields of cognition and perceptual-motor skills have used latency measures far more frequently than associative learning researchers. And it seems clear that those studying the acquisition and development of associative skills should also include the time dimension as a part of <u>their</u> measurement procedures.

Fitts and Peterson (1964) reported a reaction time experiment in which subjects learned to touch one of two targets as quickly as possible. They found that by requiring various amplitudes of movement and changing the size of the targets they could systematically alter the time to perform a correct response. However, when response latency and movement duration were measured separately, the authors found that their experimental manipulations had much greater effect on movement duration than on latency. One could conduct a similar experiment, varying the complexity of the targets, making the associative component of the task more or less difficult. Thus, even *without* measuring the separate components of total response time, it should be possible to assess the relative effects of discriminative and motor requirements of the task. Analogous forms of component analysis could be applied to the study of associative skills.

Trials Procedures and Measurement Reliability

From a series of verbal learning studies in which he measured latencies, Peterson (1965) concluded:

> It is clear that the ability of the *S* to respond correctly is a function of the length of time allowed for him to respond. In a learning experiment with fixed anticipation interval, relative frequencies of correct response will not be independent of latency (p. 167).

If the length of the anticipation interval in a trials procedure *does* affect the accuracy measure during acquisition trials, then procedures with different anticipation intervals are likely to affect trials-to-criterion. Consequently, a given accuracy criterion may not represent the same amount of learning from one procedure to another.

In a related study, White and Spiker (1960) taught subjects a motor paired-associate task which required them to choose one of six buttons in response to a given stimulus. This was an interference study in which subjects first learned an A-B association over a given number of trials, and then learned an A-C association in an assessment of negative transfer. Control subjects learned A-B, then D-C. One group had five trials on A-B, and a second group had 20 trials. In all trials, subjects were given a "lenient" 4-second anticipation interval to insure that correct responding would not be interrupted by time limits. The two groups showed no difference in accuracy on the transfer task, but measures of response latency revealed greater negative transfer for the 20-trial group than the 5-trial group. It seems likely that with shorter anticipation intervals, this study would have shown differences in accuracy. And the results seem to support Peterson's (1965) conclusion that in trials procedures, accuracy and latency measures may be misinterpreted if anticipation intervals are too short or vary from experiment to experiment.

Beck (1963) assessed the effects of allowing subjects to control the onset of each trial in a reaction-time situation. One such task required subjects to press a particular button, given a particular light. Subjects who were self-paced exhibited shorter and less variable latencies than did subjects who did not control the onset of a trial. One control group included subjects who thought they were in control of trial onset but, in fact, were not. Beck obtained similar results in experiments in which letter-sorting and pursuit-rotor tracking were the tasks.

Practical Difficulty of Measuring Latency.

One of the major reasons that associative learning researchers have not used latency measures is that it is difficult and expensive to do so. Runquist (1966), discussing instrumentation in the field of verbal learning, allowed that "the technical problems of recording latencies are considerable." Timing devices and other apparatus, sensitive enough to reliably measure single-trial latencies, are very costly. Further, subjects in such experiments require a good deal of pretraining to insure their reliability. There is little room for error when time-differences are measured in milliseconds. The combination of very small units of time and comparatively great intra-subject variability requires that experiments which measure latency be large statistical designs. This further increases the expense for research, thereby reducing the likelihood that experimenters will measure human associative learning on the time dimension.

Most associative skills do not occur in discrete, trial-by-trial fashion. Even during acquisition, such skills as reading, computation, and the control of complex apparatus generally develop in a series of continuous <u>movement cycles</u> or uninterrupted responses. It may be possible to design experiments in which such sequences of responding occur, and in which measures of performance include both accuracy and the time dimension from the very beginning of acquisition. Such procedures would eliminate the problems involved with anticipation intervals, and would also be self-paced once a performance had begun. Subsequent sections will deal with this topic in more detail.

Response Rate as a Measure of Probability

Probability and the Time Dimension

The previous section explained why the common identification of percentage correct as the measure of response probability (e.g., Judd & Glaser 1967) should be replaced with a concept of probability which includes <u>immediacy</u> as well as <u>accuracy.</u> Common sense alone would suggest that the degree to which a response is "immediately available" is an important aspect of response probability. Insofar as we observe that response latencies decrease beyond the achievement of accuracy criteria, we must agree that the probability of action increases beyond 100% correct responding.

On the other hand, response latencies are difficult to measure, and the procedures associated with their measurement do not adequately represent in vivo performance of most associative skills. Therefore, continuous measurement of responding in time, or rate of response, might provide an alternative to single-trial latencies.

Fortunately, there has been some research on the relationship between single-trial reaction times and continuous response rates. Fitts (1964) concluded from a series of experiments that "rate of responding in serial and continuous tasks agrees closely with the times obtained per response in corresponding discrete tasks" (p. 265). Heinz (1974) came to a similar conclusion from experiments comparing latency and frequency of responding in discrete-trial and continuous fixed-interval schedules of reinforcement with pigeons.

<u>Skinner's Contribution</u>

<u>Rate as the basic datum</u>. B.F. Skinner's greatest contribution to the study of behavior has perhaps been his use of response rate as a basic datum. In his first book, Skinner (1938) argued "that the rate of responding is the principal measurement of the strength of an operant. It follows that the main datum to be measured in the study of the dynamic laws of an operant is the length of time elapsing between a response and the response immediately preceding it or, in other words, the rate of responding."(p. 58) In this passage, Skinner seems to equate inter-response time with response rate, or number of responses per unit time. However, in experimental practice, these two units of measurement generally receive separate analyses. And in a later discussion, Skinner (1953b) asserted: "In the program of research to be summarized and exemplified here, probability of action has been attacked experimentally by studying the repeated appearance of an act during an appreciable interval of time." Here it is clearer that count per unit of time is the principal form for analysis, although it's reciprocal (inter-response time) has also been used to analyze a variety of phenomena.

In the glossary of a later book (Ferster & Skinner, 1957), Skinner clearly distinguished between the two data-forms. He also provided succinct definitions of two other important terms: probability of response, and strength of response. Probability of response is defined as "the probability that a response will be emitted within a specified interval, inferred from its observed frequency under comparable conditions" (p. 731). And strength of response is "sometimes used to designate probability or rate of responding" (p. 733). These statements make it difficult to distinguish between response rate, response strength, and response probability, if, indeed, a distinction is intended.

Response strength and response probability are apparently equivalent terms. But there is some question as to whether response rates always reflect response probability or strength. The question arised with contingencies of reinforcement which seem to prevent the behaver from responding freely. Differential reinforcement of low rates, for example, may suppress rate of responding. But does it also change response probability or response strength? Similar questions arise with respect to the punishment of behavior. And there seem to be two possible alternative uses of the language:

1) Response rate is a direct measurement of response probability as long as the responding is free of rate-specifying conditions. Thus, in some circumstances, rate may not reflect true response probability. Response probability is therefore an hypothetical quantity, only observable under certain conditions.

2) Response rate is always a direct measurement of probability at the time of observation. Punishment, DRL, and other contingencies which suppress rate of responding also suppress response

probability. Ratio schedules of reinforcement and interval schedules also have characteristic effects on response probability, as measured by changes in rate. Behavior never occurs in the absence of contingencies, and thus the term "freedom" is a meaningless concept in this context. In this formulation, there is no distinction between observed probability and an hypothetical strength of response.

The second of these descriptions seems preferable, from a strictly functionalist standpoint. In any case, the following discussion will treat the terms _rate_ and _probability_ as equivalent, in contrast to traditional equations of _percent_ and _probability_.

Moment-to-moment changes in probability. At the heart of Skinner's experimental analysis of behavior was the cumulative response recorder which allows the organism under study to literally "draw" its own comulative response curve, and which accentuates moment-to-moment changes in response probability. Dresslar (1892) built what may have been the first cumulative response recorder, and used it to represent the rates at which human subjects could perform finger-taps. But it was Skinner's exploitation of the device that revealed its great sensitivity to the rate of on-going behavior. The study of reinforcement schedules (Ferster & Skinner, 1957) has perhaps provided the most massive illustration of the power of this device. Countless cumulative records have revealed differences among individual organisms' patterns of behavior, and the effects of changes in the experimental environment, or history, on moment-to-moment response probabilities.

Cumulative recorders have most often been used to measure rates of such repetitive motor responses as bar-pressing, plunger-pulling, wheel-running, panel-pushing, chair-rocking, floor-pacing and key-pecking. These responses are relatively easy to transduce by electro-mechanical means. Some investigators have used voice-operated relays in conjunction with cumulative recorders to measure rates of human vocalizations. But in such applications, the measurement devices record only vocal stresses without respect to differences in topography, or verbal "content." Thus, for the most part, the experimental analysis of behavior has focussed on simple response units under the control of discriminative stimuli which remain constant across a number of responses. Seldom have functional behavior analysts attended to *rates* of complex associative skill performance in which both response topographies and discriminative stimuli may change with each response. Procedures which permit, and record the rates of, continuous responding might enhance the study of such skills as reading, writing, and computation.

Two Studies Involving Continuous Recording of Associative Skills

A program of classroom research based on average response rates led to an interest in obtaining cumulative records of associative skill performance. What could such moment-to-moment recordings reveal that is not evident from day-to-day changes in average rates of skill performance? The following two studies represent responses to that research question.

I: Some Relationships Between Component and Composite Skills

Subjects. Subjects were two male graduate students. One, of whom had assisted in research to be reported in Chapter V, was familiar with the stimulus material. The other was not.

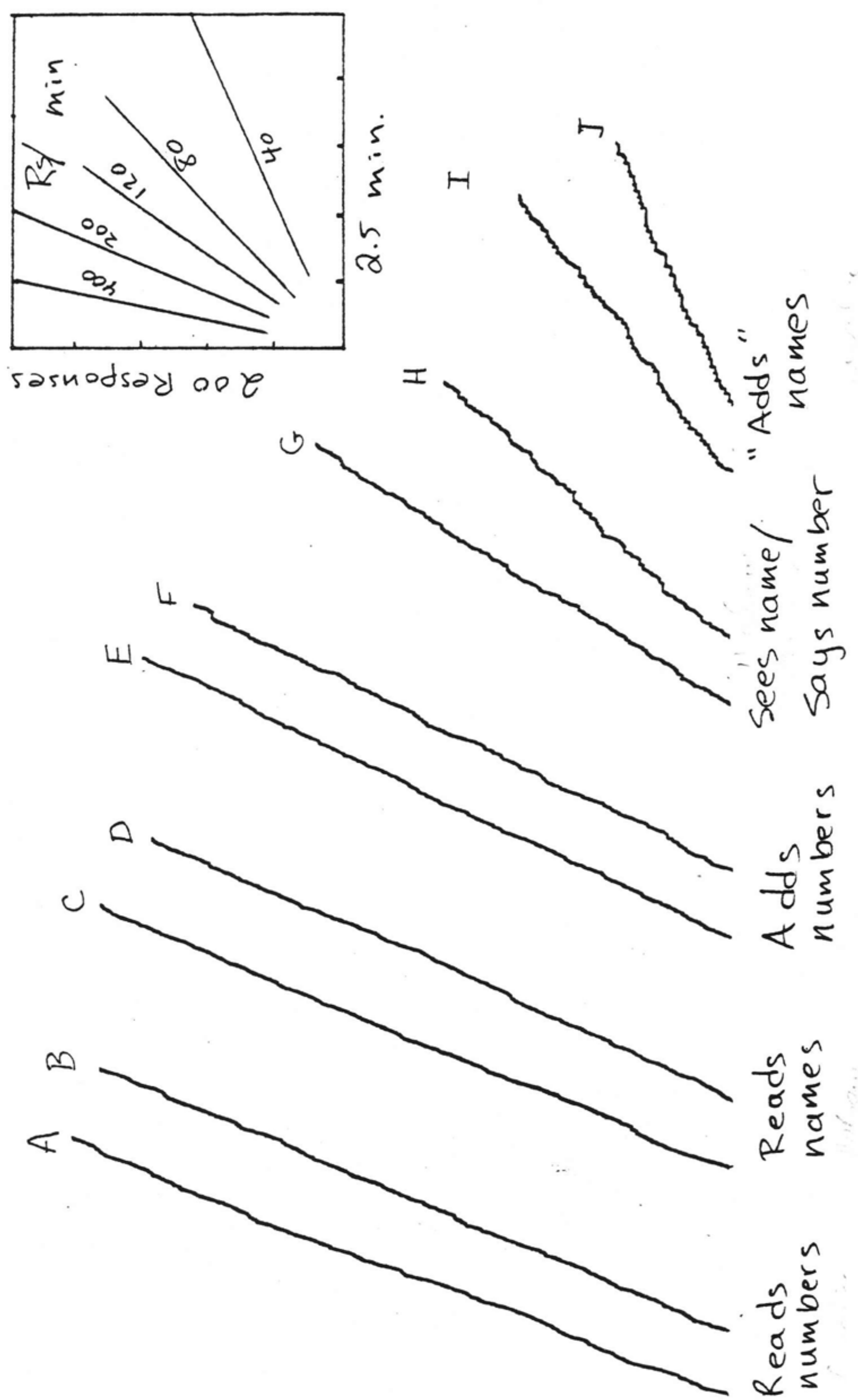
Equipment and materials. Apparatus included standard electromechanical programming equipment, timers, counters, and a cumulative recorder. The acquisition phase also required two Kodak 760H carousel projectors and a rear projection screen. A voice-operated relay and microphone transduced the subjects' spoken responses. Materials included a number of worksheets of stimulus arrays from which subjects read numbers, words, addition problems, etc. All sheets were standard typing paper, and stimuli were arrayed in columns and rows, 1 inch center-to-center.

Procedure. In the acquisition phase, the naive subject learned to say a given numeral (0-9) in response to a printed word (Anglicized names of Hebrew characters). This was a standard paired-associate anticipation procedure, and used the two slide projectors to present stimuli and responses on a screen. The anticipation interval was 3 seconds, and the pair of items appeared together for 2 seconds. The subject learned a total of 10 such pairs to a criterion of three times through the entire set with no errors. The other subject learned the name-number associations by assisting in other phases of the research.

After acquisition, a series of 2-minute tests required subjects to make oral responses as quickly as possible. The tests included: reading numerals, reading Anglicized Hebrew names, and saying "sums", in English, in response to pairs of Anglicized Hebrew names arranged as vertical addition problems on a worksheet. The last task involves mediated transfer of training by means of pairings learned in the acquisition phase.

Results and discussion. All performances were essentially 100% accurate. All tests were recorded on a cumulative recorder, timers, and counters, via a voice-operated relay. Figure 1 presents cumulative records from

Figure 1



200 Responses
2.5 min.
Rs/ min
400
200
120
80
40
A
B
C
D
E
F
G
H
I
J
Reads numbers
Reads names
Adds numbers
Sees name/ Says number
"Adds" names

one set of tests. On cumulative response records, slopes are proportional to response-rate: changes in slope indicate changes in rate from moment to moment. The records appear in pairs, each pair corresponding to a particular task as performed by both subjects. The experienced subject's record is the first in each pair. Record _A_ represents a number-reading performance of approximately 198 per minute, on the average. Record _J_, the slowest, represents "adding" names at about 36 responses per minute on the average.

Several aspects of these records deserve mention. The most obvious is their regularity. Especially on the first three tasks, the pace of these performances was extremely regular over the 2-minute periods. The eveness of the records suggests that average response rates over appreciable periods of time are comparatively undistorted representations of such performances. Some of the irregularities in these records reflect pauses for breath at regular intervals.

A second important aspect is the orderly decrease in rates from the most familiar skill, reading numbers, to the most difficult one, "adding" Anglicized Hebrew names. Reading numbers is a component of adding numbers. And even highly proficient adult subjects, for whom simple addition is "automatic," read numbers more quickly than they add numbers. The component-composite relationship is even more apparent in the comparison between the first three tasks and the last two. In particular, "adding" Anglicized Hebrew names requires subjects to read the names, covertly respond with numerals, and add those numerals. The subject who is less proficient at saying numbers in response to names also shows a lower rate in the mediated transfer task of "adding" names.

II: Effects of Distracting Auditory Stimuli on Performance Rates

LaBerge and Samuels (1974) presented an information processing theory of what they call <u>automaticity</u> in the development of reading proficiency. Arguing that less attention is required for skills that have become automatic, they suggested that a distracting or surprising stimulus would interfere more with a nonautomatic skill than with an automatic skill. Their basic procedure involved trials in which subjects indicated whether two letters or characters matched by pressing a button, or not. On most trials, letters appeared one at a time. The first appeared, and then when the second appeared the subject responded as quickly as possible if the letters were the same. Occasionally, the first letter was followed by an entirely different <u>pair</u> of letters. At that time subjects were supposed to respond, or not, depending on whether the two letters were the same. When the pairs in these "surprise" trials were well-known alphanumeric characters, subjects responded with relatively short latencies. However, when unfamiliar visual characters appeared, latencies were somewhat more protracted. Laberge and Samuels (1974) used these differences in latency to define automaticity, and they constructed a multi-level model of coders, processors, and attentional mechanisms to account for the effect.

Many of the practical implications of this research are consistent with conclusions drawn from a more functionalist approach, and will be part of the discussion in Chapter III. But the possibility of constructing a related paradigm, using freely emitted behavior and response rate measurement, led to the following study.

<u>Subjects</u>. Subjects were the same as in the first study.

<u>Equipment and Materials</u>. In addition to the equipment and materials for the first study, this experiment required a tape recorder and earphones.

<u>Procedure</u>: Subjects took a set of tests identical to those in the

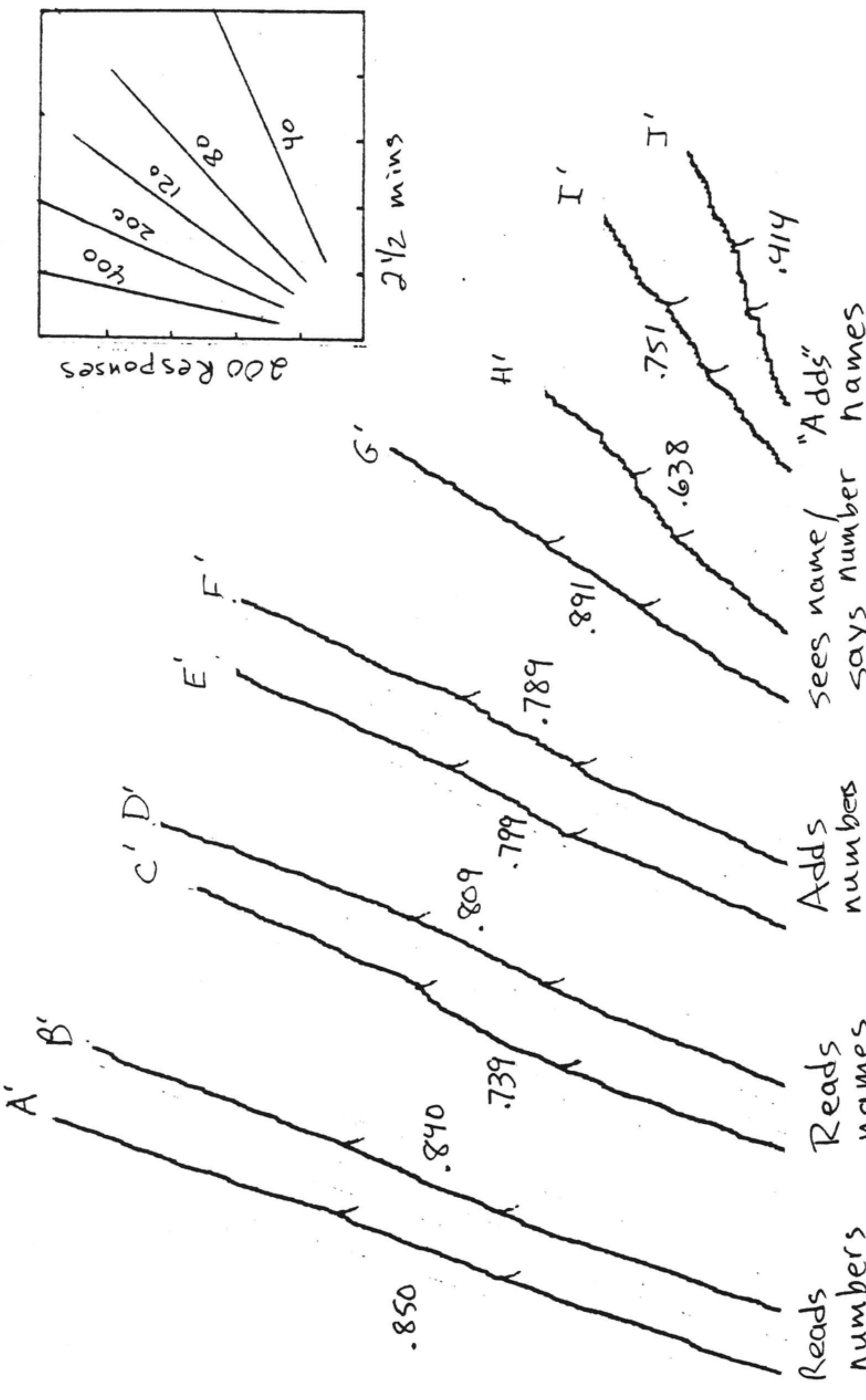
first experiment. However, during the middle 30 seconds of the 2-minute test periods, they listened to a tape recording of someone reading random numbers.

Results and discussion. Figure 2 shows cumulative records, arranged in the same manner as the records in Figure 1. The two deflections of the pen on each record indicate the beginning and end of the 30-second distraction period during which subjects heard a recording of number-naming through earphones, while continuing to perform the task. The numbers next to each of these periods on the records represent ratios of average response rate during the distraction period to average rate before and after the period. They are suppression ratios, as used by Estes and Skinner (1941) in the study of conditioned suppression of an ongoing operant baseline.

These records display a number of regularities. First, the records display performances outside of the distraction period that are virtually identical to those observed in the first study (Figure 1). In general, Figure 2 shows the same relationships among skills and between subjects as does Figure 1. However, the addition of the distraction period reveals another set of relationships. The distraction period had comparatively less effect on performance of the first three skills, which are highly developed for both subjects, than on the last two. And on the last two skills, there was greater suppression for the subject who had recently acquired them, than for the more proficient subject. On the "adds names" performance, the less proficient subject responded only eight times during the distraction period.

Repetitions of the tests demonstrated that the procedure produces reliable results. Tests conducted later the same day yielded suppression

Figure 2



200 Responses
2½ mins
400
200
120
80
40
A'
B'
C'
D'
E'
F'
G'
H'
I'
J'
.850
.840
.739
.809
.799
.789
.891
.638
.751
.414
Reads numbers
Reads names
Adds numbers
sees name/ says number
"Adds" names

ratios varying from those in the first series by no more than .056 and by as little as .005. Because of the apparent relationship between skill proficiency and suppression, we would expect the ratios to grow as a function of practice on a given skill. Thus, repeated testing would probably not produce equivalent results, but rather, an orderly increase in the ratios as the skills became more fluent.

General Discussion

These results are important because they represent a new methodology. They show that continuous recordings of associative skill performances can reveal important relationships among skills, and between various independent variables and the performance of those skills. Recordings which show moment-to-moment changes in response rate illustrate the effects of practice on the regularity of responding, and the effects of distraction on more or less proficient skills.

This method should make it possible to discover response rate criteria for skill development. It would seem that one measure of associative skill proficiency should be the degree to which a skill is susceptible to distracting environmental events. The procedure used in Experiment II should allow us to determine at what rate a specific skill must occur in order to become relatively impervious to distraction. We need only study the relationships between undistracted rates and suppression ratios, across tasks, and histories of practice. This method allows for a functional definition of "automaticity" without recourse to the complex hypothetical constructs posited by LaBerge and Samuels (1974).

Terrace (1966) defined <u>stimulus control</u> in terms of the degree to which change in an antecedent stimulus determines the probability of responding. In Experiment II, there was a change in response probability during presentation of the distracting stimulus. This effect can be attributed to conflicting degrees of stimulus control exerted by the task stimuli and by the distracting stimuli. In the less proficient tasks, the task stimuli exert comparatively less control than in the more proficient tasks. Thus, in less proficient tasks, introduction of competing stimulus control has a comparatively large effect on probability of responding, and produces a large degree of suppression in task response rate. Discussing the concept of _attention_, Skinner (1953) said "When someone is paying attention he is under special control of a stimulus" (p. 123). We may thus describe the relationship between skill proficiency and the ability to maintain attention to task in functional terms, relating rate of responding to degree of suppression due to a distracting stimulus.

Response Rate as a Measure of Skill Proficiency

Skill Proficiency and Probability

The results and discussion of the previous section used the terms _probability_ and _proficiency_ interchangeably. If response rate is the appropriate measure of probability, then it is also a measure of proficiency. Probability has traditionally been equated with percentage correct, and most educators measure proficiency on the percentage correct scale. But it is clear from the literature review in earlier portions of this chapter that measures of proficiency must go beyond accuracy criteria.

In referring to _mastery_, a high level of skill proficiency, Gagne (1974) described skills which had been mastered as being "immediately accessible" for application to more complex skills. Elsewhere (Gagne and Briggs, 1974), he described proficiency as the ability to "perform with perfect confidence," not just "sometimes." In a related vein, Skinner (1953a), who measures probability as response rate, observed that the term knowledge is sometimes used "to represent simply the probability of skilled behavior." Improvements in skill or knowledge, then, are changes in response probability, And it is appropriate to measure response probability, in freely emitted sequences of behavior, as rate of response.

Certain skills (e.g., speech) are best performed at rates below the maximum possible, as well as above a specified minimum. This is usually because of undesirable changes in response topography at extremely high rates. In such cases, proficiency is not isomorphic with response rate beyond a specified point. But even in those cases, rate measurement allows proficiency to be assessed in a manner beyond the capabilities of a simple accuracy or quality measurement. Moreover, most associative skills do not have upper limits of proficiency on the rate scale.

Procedures Which Impose Proficiency Limits

Most associative skills involve continuous sequences of movement cycles for appreciable periods of time. However, most procedures for the study of associative learning, many perceptual-motor skill paradigms, and many educational procedures (especially in special education) allow only one response to occur at a time. Or, even if they do provide opportunities for multiple responses, the opportunities are limited. Such procedures impose ceilings, both on our ability to measure proficiency and on the subject's or student's ability to develop proficiency. Such procedures have been associated, historically, with the general lack of attention to the time dimension, as discussed earlier.

Particularly in applied contexts, these procedures handicap students. And in research situations, they obviously prevent us from examining the dynamics of ongoing performances and from measuring learning on the response rate scale. A few examples should clarify this point.

In a sight vocabulary training program for severely retarded adolescents to be reported in Chapter IV, students initially learned to read printed words from flashcards presented by the teacher one at a time. Because it is impossible to present more than 10 to 15 cards per minute, the procedure prevented students from increasing their reading speed. (Normal adults read aloud at a rate of 200 to 300 words per minute.) Most teachers of this population would train students to an accuracy criterion with such a procedure, and then move on to a new set of words, never allowing students to develop greater reading proficiency. However, in this case the teacher constructed worksheets from which students could read words as quickly as possible, limited only by their own proficiency. On the first day of this new procedure, a typical student read 35 words in a minute. The procedure allowed an unconstrained assessment of his reading proficiency.

In a paired-associate learning experiment (to be reported in Chapter V), subjects learned to name Hebrew characters in an anticipation procedure. The anticipation interval was 3 seconds, and the character and its name appeared together for 2 seconds. The intertrial interval was approximately .8 seconds. This procedure thus constrained subjects to a rate of approximately 10 responses per minute. Yet a typical subject was able to name 62 characters in one minute from a worksheet, immediately after meeting a 100% accuracy criterion for a set of ten characters. Subsequently, some subjects were able to name as many as 120 characters per minute during the first practice session.

These data highlight the need for designing procedures that do not constrain performance. An immediate effect of the change from percentage correct to response-rate measurement is critical evaluation of procedure-imposed ceilings on performance.

Sensitivity of Response Rate Measurement

Response rate data on human performance are already available from a number of sources. Rate-based tests of reading and typing are common, and there are many rate data on fine motor skills. Chapter III reviews an approach to education based on daily measurement of response rates. But rate measurement has not been incorporated into a systematic framework of research issues, methods, and measurement procedures for the study of associative skills. Such systematic research might be of great practical value for instructional design, revealing important functional relations.

Some of the human response rate data which already exist suggest that the rate dimension is uniquely sensitive to important characteristics of human behavior. What follows is a sampling of these findings, including some data from my own work with teachers.

Response Rate and Measures of Intelligence

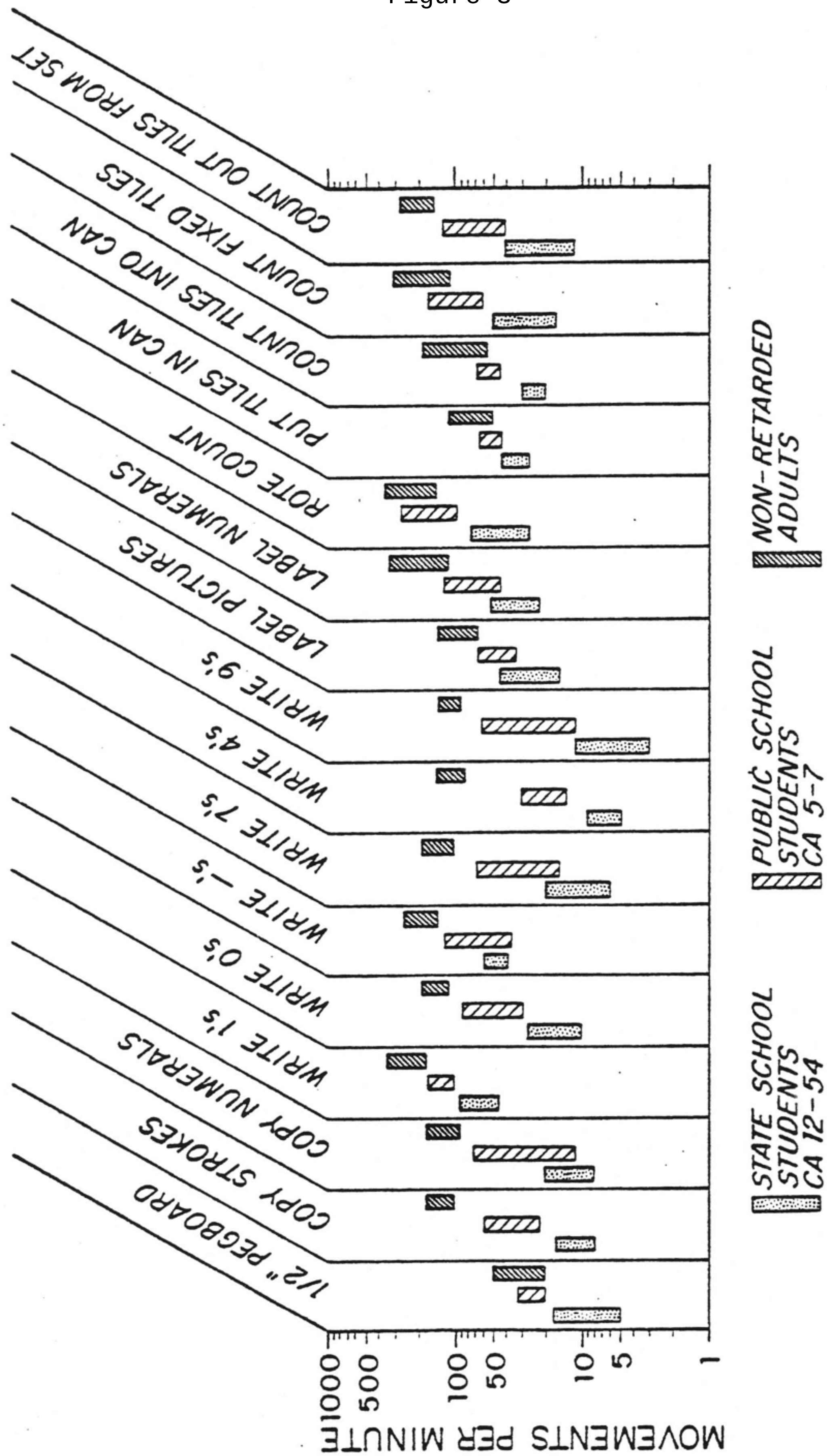
Barrett (1965: Barrett & Lindsley, 1962) collected more than 38,000 hours of free operant conditioning data from more than 100 institutionalized mentally retarded subjects. This work followed upon that of Lindsley (1962) with psychotics. It involved subjects, often for several hundred experimental sessions, in the operation of manipulanda under various discrimination contingencies for a wide range of reinforcing consequences. Subjects were exposed to these contingencies in 6' x 6' sound-attenuated operant conditioning chambers. Automated apparatus recorded multiple behaviors, including vocalizations, pacing, rocking in a chair, pulling and pushing small wall-mounted plungers, and touching panels.

Barrett's subjects ranged from mildly to severely and profoundly retarded, as defined by standard psychometric evaluations. Her focus was on functional definition of deficits in terms of response to various contingencies of reinforcement. She (Barrett, Note 1) reported that 70% of the severely and profoundly retarded participants operated the apparatus at a median rate below 20 per minute, while 64% of the less retarded participants responded at median rates of 20 per minute or above. Rate of responding was also positively correlated with age, and with age at admission to the institution. On the other hand, there was no difference between the two psychometrically defined groups in the degrees of behavioral efficiency which they eventually achieved. Although the more retarded subjects required longer to reach their best levels of operant discrimination and differentiation, they did not ultimately differ in terms of what they were able to learn. Thus response rate, but not accuracy or efficiency of performance, was a correlate of psychometric level. This finding, if replicated in the performance of associative skills, has important implications for instructional design in special education.

<u>Rates of Skill Performance, Age, and IQ</u>

Figure 3 shows results from a pilot study reported by Barrett (1979). Four individuals from each of three populations had repeated opportunities to perform 16 simple skills. The populations were: normal adults, young public school students, and institutionalized retarded students. At each opportunity to perform, the subjects completed as much work as possible in a 30-second period. All subjects, including the retarded, performed at essentially 100% accuracy. Thus it would be impossible to distinguish among the performances on a percentage correct scale, since they reflect high degrees of <u>overlearning</u>. However, on a response rate scale, all the public school students exhibited greater proficiency on 11 of the 16 tasks than the retarded subjects. These data illustrate a variety of re-

Figure 3



Frequency comparisons on some components and prerequisites of elementary skills.
(Based on an unpublished pilot study conducted by Frances George and Deborah Pease)

lations among skills, but the major point is that the rate dimension was sensitive to differences that the percentage scale would have obliterated.

Identification of Learning Problems and Gifted Children

Spring (1975) tested 44 first grade children on a digit-naming task for 1 minute and administered the Cooperative Primary Reading Test to each child. He found that the digit-naming rates differentiated with about 90% accuracy between poor and average readers, as defined by the reading test. In another study (Spring, 1976), he found that "dyslexic" students in a regular public school were slower than other students on a number-reading task, but performed within normal ranges on picture-naming and color-naming tasks.

A massive project in Washington State (Child Service Demonstration Programs, Note 2), collected approximately 150,000 rate samples on nearly 3,000 skills from 17,996 students in three school districts. Teachers collected 7 to 10 days of data per student per skill to determine median rates as well as trends. Samples were 30 seconds or 1 minute. Each student took tests on a relatively small number of skills. Project staff defined problem learners as students with less than half the median class rate or median class acceleration (change per week in rate) on a given skill. This procedure identified more than 70% of the students whom more costly procedures had diagnosed as having learning problems. It also identified some students not found by the more elaborate procedures. Duncan (1969) recorded tapping, walking, reading, and calculating rates for 46 gifted and 30 regular school children. She also administered standard achievement tests. The gifted as a group performed significantly faster than other students.

Rates Change with Age

Edwards and Edwards (1969) reported rates of fetal movements, as counted by pregnant women. They found an orderly increase from about one every 100 minutes, on the average, at 5 months of pregnancy, to nearly 1 per minute several weeks before birth. There are few published data on post-natal development of movement rates, although this would seem to be a rich field of study. Roberts, Bondy, Mira, and Cairns (197) reported data from two normal infants. Their patterns of development were striking similar. The authors observed the infants for 5 minutes a day and counted three response topographies: lifting the chin from prone; lifting head and chest (defined by visibility of a piece of tape between the nipples); lifting self on hands and knees. They observed one infant from 12 to 32 weeks, and the other from 9 to 34 weeks. The frequency of chin-lifts decreased from the beginning, as it was replaced by the other movements. Lifting on hands and knees began after 8 to 10 weeks, and accelerated rapidly. Chest-lifts accelerated only slightly and, in one infant, began to decelerate as lifting self on hands and knees replaced it. The two children displayed remarkably similar rates of change on all three movements. Although neither these gross body movements nor fetal kicks are associative skills, they are obviously important prerequisites. And the order of the data on the rate scale is remarkable.

Connolly, Brown, and Bassett (1968) tested 60 children, aged 6, 8, and 10, on a simple perceptual-motor task. They required the subjects to repeatedly dot two paper targets in alternation with a pencil, for 5 seconds at a time. Girls of a given age performed faster than boys, and older children were faster than younger children. However, there were no sex or age differences in accuracy.

Doehring (1976) and Denckla and Rudel (1974) tested elementary school children on a variety of skills related to reading. Skills included naming objects, colors, numbers, animal pictures, letters, syllables, and words in various categories. The two studies produced similar results, showing orderly increases in all skills from kindergarten through 11th grade.

Rate as an Index of Skill Difficulty

Fitts (1954) tested subjects on three psychomotor tasks: placing discs on pins, putting pegs in pegboard, and alternatively tapping targets with the index finger. He found that rates increased uniformly as movement amplitude was decreased and tolerance limits (target sizes) extended. These data, as well as Barrett's (Figure 3), illustrate how tasks may be scaled on a rate dimension.

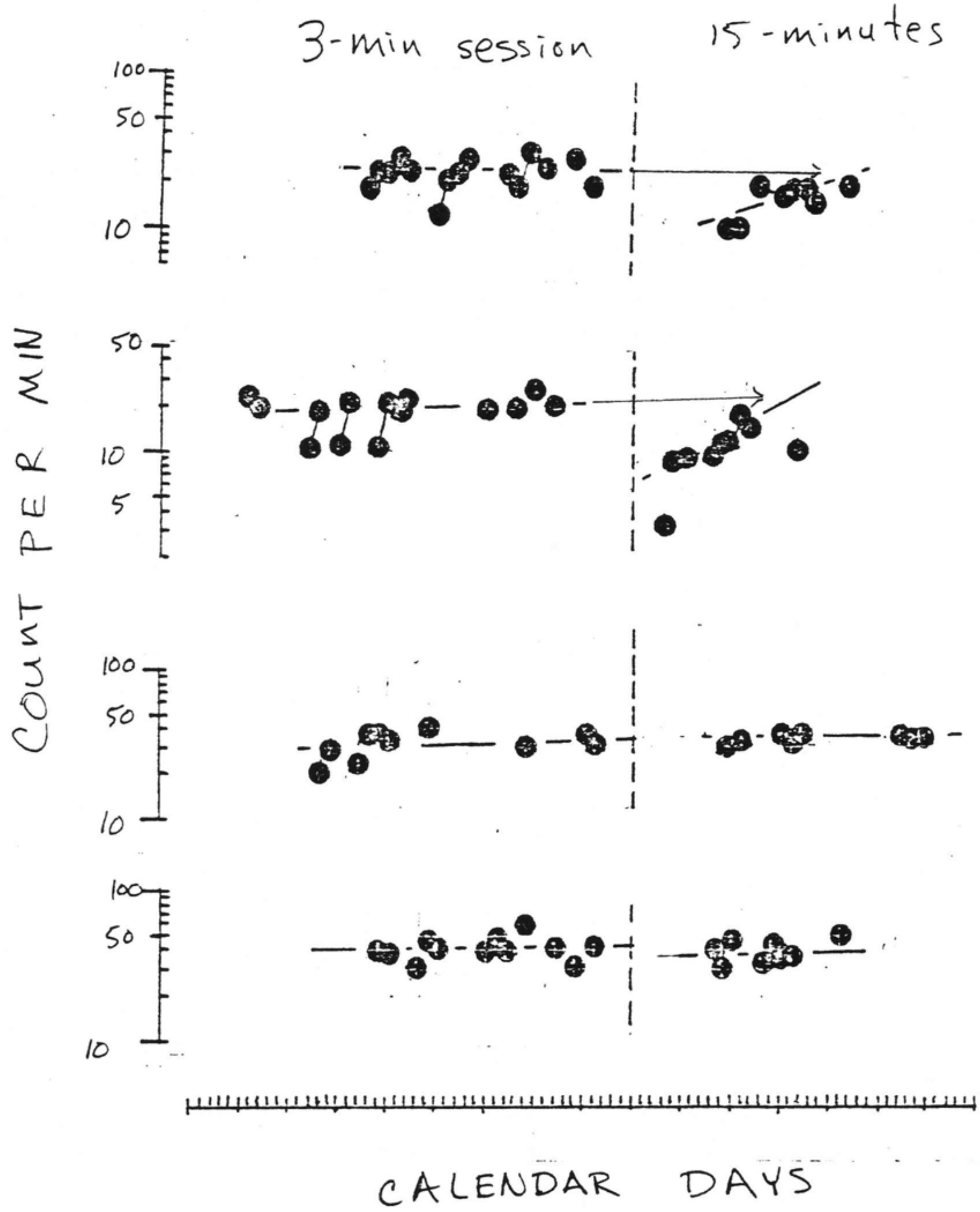
Rate as a Measure of Professional Discourse

In a self-paced college psychology course, Johnston and Pennypacker (1971) recorded students' rates of oral responding to questions about psychology, arguing that participation in a discipline at a professional level requires ability to speak fluently on relevant topics.

Rate, Endurance, and Attention

The rate at which an individual can perform a given skill for a short period of time may predict how well that person will be able to _endure_, or maintain that rate for a longer period of time. Figure 4 illustrates this relationship. Subjects were four severely retarded adolescents attending classes in the Behavior Prosthesis Laboratory at the Fernald State School (Waltham, MA.). They were required to count out a number of small items to match the numeral written in each compartment of an egg carton. All subjects performed at a high level of accuracy, and only rates of correctly counted items appear in the figure. A pharmacist, highly proficient in counting pills, performed the task at about 80 items per minute. For several weeks all subjects practiced this skill for three minutes per day, and their daily

Figure 4


3-min session
15-minutes
COUNT PER MIN
100
50
10
50
10
5
100
50
10
100
50
10
CALENDAR DAYS

rates appear in the first half of each graph. Two subjects averaged about 20 to 25 items per minute, while the other two counted 30 to 50 per minute. Dotted vertical lines mark the point at which teachers changed the practice period from 3 to 15 minutes. The two slower students slowed substantial decrements in average performance rates, while the more proficient students maintained their earlier average rates. Although they are by no means conclusive, these data are remarkably orderly.

One may think of <u>endurance</u> in terms of attention to task, as well as in terms of resistance to fatigue. In fact, the two phenomena seem to be related in many classroom tasks. We often observe that students who are able to perform at relatively high rates are comparatively resistant to distraction, "daydreaming," and the like, while less proficient students seem more likely to engage in alternative behaviors. Often, by returning the less proficient students to component tasks which they can perform at higher rates, we observe greater attention to task as well as more rapid progress. This finding may be related to the phenomenon revealed by the cumulative records shown earlier. Just as certain levels of proficiency insure resistance to distraction, they also allow individuals to maintain performance for longer periods of time. These phenomena have significant implications for instructional design, and deserve additional systematic research.

Summary and Conclusions

By ignoring the time dimension, percentage correct measurement imposes a ceiling on our ability to detect and assess learning beyond mere acquisition. Response latency measures, which continue to change after subjects reach acquisition criteria, are sensitive to so-called overlearning. But latency measurements are costly and difficult to obtain reliably.

Skinner introduced response rate as a measure of probability. We can measure associative skills on the rate dimension and, by using a cumulative recorder, detect moment-to-moment changes in response probability.

The term proficiency implies a high degree of response probability under appropriate conditions. We can measure associative skill proficiency in terms of response rates. To do so, we must remove procedural constraints on freely emitted behavior.

Response rate is a sensitive index or correlate of age, intelligence score, learning disability or "giftedness," skill difficulty, fluency of professional discourse, and endurance or attention to task.

Chapter III: Measuring Response Rates in the Classroom

Rationale

From Laboratory to Classroom

After more than a decade of human operant conditioning research in the laboratory, Lindsley decided to take the rationale and measurement technology of an experimental analysis of behavior into classrooms. His first paper in the field of education (Lindsley, 1964) began:

> Children are not retarded. Only their behavior in average environments is sometimes retarded. In fact, it is modern science's ability to design suitable environments for these children that is retarded. (p. 62)

Laboratory analyses of retarded behavior (Barrett & Lindsley, 1962), convinced Lindsley that the methodology of behavior analysis could be applied in classrooms.

A primary ingredient of an experimental analysis of behavior is the distinction between operations and functions. Operational definition describes a series of events, or operations, which the experimenter arranges in the hope of affecting the behavior under observation. Through systematic manipulation of operationally defined events, in conjunction with behavioral measurement, the scientist may discover functional relationships between an individual's behavior and environmental events (Skinner, 1953a, p. 35). Because many educators (and even behavioral researchers) have failed to distinguish between operations and functions, Lindsley (1964) expressed "the four-component operant behavioral equation" in two different forms. The form corresponding to operational definition is:

$$E^A \quad - \quad M \quad - \quad A \quad - \quad E^S.$$

In temporal sequence, an antecedent event (E^A) occurs before a movement (M) on the part of the individual. The movement is then followed by a <u>subsequent event</u> (E^S) with a particular <u>arrangement</u> (A).

In an experimental analysis, the scientist may verify the <u>functions</u> of these events. Accordingly, the second form of the equation is:

$$S \quad - \quad R \quad - \quad K \quad - \quad C.$$

Antecedent events may function as <u>stimuli</u> (S), controlling probability of response. Movements may function as <u>responses</u> (R),operating on the environment to produce effects. Subsequent events may function as accelerating or decelerating <u>consequences</u> (C), if the manner in which they are arranged to follow movements functions as a <u>contingency</u> (K) for the behaver. With this emphasis on functional analysis, Lindsley proposed an approach to education, based on measurement. Lindsley used the term <u>movement cycle</u>, to describe potential responses. It may be coincidental that this term also appears in perceptual-motor learning literature (e.g., Fitts, 1964). There was a precedent in laboratory operant conditioning for studying repeatable movements which could be counted by automatic apparatus. Lindsley and his associates argued that such movement cycles, rather than bodily positions or orientations (e.g., "keeps hands in lap"), constituted the most important targets for applied behavior analysis.

Unlike other behavioral psychologists who have moved from basic research into application, Lindsley did not seek recipes for behavior change. In fact, this may be the major difference between his "precision teaching" approach and other forms of applied behavior analysis. Emphasizing the individual organism, as Skinner did (1956), Lindsley reasoned that effective education must take individual differences into account. He argued that instead of recipes, educators need a measurement system sensitive to the effectiveness of individualized procedures (Lindsley, 1972). With such an approach, teachers would be able to discover which procedures are functional for which individuals, and perhaps find generally effective procedures as well. The method, like that of an experimental analysis of behavior, was to be strictly inductive.

Measurement as a Communication Device

In an analysis of behavior, measurement functions as a communication device in two ways. First, it allows the behaver to communicate to the behavior analyst (Barrett, 1977). Skinner (1948, p. 289) argued that the "organism is always right." Or as Lindsley (1972) put it, in the educational context, "The child knows best." The point is that measurement reveals the functions for the behaver of events arranged by the educator or scientist. It allows the scientist or educator to decide whether or not specific procedures are effective, or functional. In addition, measurement functions as a medium of communication among behavior analysts, all the more so with standardized forms of measurement. The cumulative record, for example, provides for extremely efficient communication among experimental behavior analysts. One need only know the step-size and paper-speed to be able to interpret a cumulative record. Because of his long experience with cumulative records, Lindsley (1972) decided to design a standard behavior chart for educators and other behavior-change agents.

The Standard Behavior Chart

Although it is currently in its ninth revision, the Standard Behavior Chart has remained essentially the same since its invention in 1965. It allows for a standardized display of any countable behavior, thus providing a means of efficient communication and a framework for viewing a range of research questions and behavioral dimensions. The following is a summary of its basic display characteristics and summary quantifiers. For a more thorough discussion, see Pennypacker, Koenig, and Lindsley (1972), White and Haring (1976), or White and Liberty (1976).

Calendar Base and Synchronization

Because the Standard Behavior Chart is based on calendar days, not sessions, one can observe the effects of intersession intervals. The Daily Behavior Chart, spanning 140 days, is especially valuable in school applications. (There are also weekly, monthly, and yearly versions.) In North America, it is conventional to begin the first chart of the year on the Sunday before Labor Day (White and Haring, 1976). Especially within a given program, school, or research project, such calender synchronization is useful for examining patterns of behavior across behavers, or within behavers across behaviors. Often it is possible to discover behavioral covariation simply by superimposing two or more charts on a light-box.

Count Per Minute Ordinate

The Daily Behavior Chart reduces all data to count per minute. Lindsley adopted the response rate measure which had been so useful in laboratory operant conditioning (Skinner, 1950, 1953c, 1956).

Durations, Latencies, Record Floors

It is also possible to record periods of time on the Standard Chart. Since rate is the reciprocal of duration, time measurements may be charted as the rate corresponding to one response in the specified amount of time. For example, a duration of 12 seconds would appear as a count of 5 per minute. A duration of 10 minutes would appear as .10 per minute. A latency of 250 milliseconds would appear as 240/minute. The record floor or counting period floor represents the length of the recording period. It is important to be able to graph this dimension along with counts per minute. The endurance data in Chapter II provide one example of how these two measurements might be functionally related. It is conventional to chart the record floor as a small dash for each day on which rate, latency, OT duration data, are also charted. The record floor enables the charter to make judgments about measurement procedures. A time period may be too short to permit observation of any instances of comparatively low-frequency behaviors. Systematic variation of the counting period allows calibration of the measurement procedure. Comparisons of rates obtained across a range of counting periods yield information about appropriate observation or practice durations.

When an observation yields a count of zero for a given counting period, the charter puts a question-mark just below the record floor. This convention accentuates the fact that a longer observation period might have yielded a count greater than zero.

A zero-count for 5 minutes is different from a zero-count for one hour in that a response which occurs, say, 10 times per hour might not occur at all during a 5-minute observation. This is similar to the property of percentage correct measurement whereby the number of opportunities to respond determines the size of possible percentage increments. It is impossible to achieve a score of 25% correct when there are 10 opportunities to respond. White and Haring (1976) have called the increment-size on the percent scale a percentage record floor.

Logarithmic Transformation on the Ordinate

Whereas the calendar base is an equal-interval scale, the ordinate of the Standard Behavior Chart is logarithmic or equal-ratio. This characteristic of the chart has several advantages.

Wide range. The ordinate spans a range from 1000 movements per minute to .000695 movements per minute, which is one response per day. Although such behaviors as silent reading or masterful banjo-picking may occur at a higher rate than 1000 per minute, most skills are within the 6-cycle range. Thus, virtually any human response rate may appear on the same chart, for convenience as well as for comparison.

Accuracy Ratios. Precision teachers chart rates of correct responses, errors, prompted responses and other categories of skilled performance as separate points on the same day-line. Rate measurement allows independent analysis of the probability of each category of response. This is not possible with percentage measurement.

The logarithmic ordinate of the Standard Chart makes it possible to express accuracy as a ratio. Thus, rate of correct responses may be a multiple (e.g., x 2.0) or a fraction (e.g., ÷ 1.5) of the error rate. Such accuracy ratios quantify accuracy, independent of rate (Pennypacker, et. al., 1972). Standard Chart represents a given accuracy ratio as a vertical distance between the points for error rate and rate of correct responses on a particular day. For example, a x 3.0 accuracy ratio (75% correct) appears as the distance between 1.0 errors per minute and 3.0 correct responses per minute, 5.0 and 15.0, or any other multiple of x 3.0. These distances are the same on the logarithmic ordinate.

If there is a zero count for a category of response, the value of the record floor is used in place of zero in computing the accuracy ratio. This is reasonable since a longer observation period with a count of no errors, for example, verifies a greater ratio between correct responses and errors than does a shorter period. The effect of this convention is to provide a scale of accuracy which goes beyond 100% correct. For even with a count of zero errors, it is possible either to increase the rate of correct responding or to increase the duration of errorless performance. Either change would increase the accuracy ratio.

Trends or progress lines. The logarithmic ordinate allows for the expression of trends as ratios. A rate of behavior may multiply (accelerate) or divide (decelerate) to over time. Trends are expressed as 'celerations on the Standard Chart, and a given trend (e.g., x1.50) appears with the same slope anywhere on the chart, irrespective of its starting-point.

In general, users of the chart apply the procedures of resistant-fit exploratory data analysis (Tukey, 1977). That is, they use medians rather than means as indicators of central tendency, quarters rather than standard deviations as indicators of dispersion, and so forth. These techniques moderate the effects of unusually high or low performances on the overall analysis, and they enable charters to use simple mechanical methods to summarize data. Among these techniques is the quarter-intersect method of drawing 'celeration lines (White, Note 3: Pennypacker, et al., 1972: White and Haring, 1976). Quarter-intersect trend lines, when assigned 'celeration values (e.g. x3.0, ÷1.50) summarize the effects of procedures, and individuals' behavior change over periods of time. It is possible to summarize a given project, phase, or course of behavior change by stating the starting rate, 'celeration, and ending rate.

Variability as a ratio. Having obtained a 'celeration line, it is then possible to draw an envelope of two lines, parallel to the 'celeration line, which encompasses all or any desired percentage of the data points. Koenig (Note 4) and White (Note 5) analyzed the effects of trimming various percentages of data points from the envelope. Their findings were similar to the general conclusion (e.g., Wainer, 1976) that trimming to 50-75% provides a stable, yet still representative value. Pennypacker et al. (1972) reported that on the Standard Behavior Chart human behavior frequencies show a high degree of homogeneity of variance. They wrote:

> This observation, combined with the fact that increases or decreases in behavior frequencies are accurately described by a straight celer-

> ation line, allows us to make surprisingly accurate predictions using these simple chart procedures. Charting behavior frequencies on anything other than a ratio scale would require us to resort to extremely complex mathematical operations for making these same predictions. The extensive historical use of non-ratio charts in psychology and education probably accounts for the fact that very little hitherto has been accomplished in the area of predicting human behavior, despite widespread agreement that such prediction is both possible and desirable (p. 100).

The envolope of bounce can be summarized as the ratio of the upper to the lower boundaries. We find, for example, that the normal adult population exhibits a range of between x1.50 and x2.50 on most well-practiced associative skills, whereas handicapped populations may vary a great deal more. Haring (Note 6) used 50% envelopes of bounce to distinguish between severely handicapped students with "compliance problems" and those who did not exhibit such irregular response to instructions from day to day in instructional programs.

Change in rate as a ratio. A given change in procedure might produce an incremental change in frequency, or rate. For example, the change from trials procedures to worksheet performance in the reading program described in Chapter II was a frequency multiplier of approximately x3.0.

Change in trend as a ratio. A given change in procedure may alter the celeration. A 'celeration multiplier is an index of this change. If the two 'celerations have the same sign (x or ÷), one divides the larger by the smaller and gives the quotient the sign of the change ("x" for increase toward positive slope, "÷" for movement toward negative slope). If the signs of the two'celerations differ, one multiplies the two and gives the product the sign of the change (Pennypacker et al., 1972).

Change in accuracy as a ratio. A given procedure may produce a change in accuracy. An improvement index is a 'celeration multiplier of the correct rate celeration and the error celeration (Pennypacker et al., 1972).

In summary, the logarithmic ordinate of the Standard Behavior Chart offers a substantial number of practical advantages. The chart is a tool of great versatility and efficiency for the analysis of behavior and for the communication of results.

Is the Semi-logarithmic Chart Best?

Scientists transform numerical data when the transformation provides advantages for graphic or statistical analysis, or when they believe that the chosen transformation represents a more accurate model of natural phenomena. In the case of the Standard Behavior Chart, it is clear that the logarithmic ordinate and linear abscissa offer important practical and analytic advantages. Nonetheless, it is still possible to question whether the semi-logarithmic display accurately models changes in response rate over time.

There are a variety of possible objections to using a semi-log chart rather than one with a linear ordinate. The most obvious is that the ratio chart obscures small differences between comparatively high response rates. For example, whereas, the difference between 10 and 20 per minute is clear on the chart, it is nearly impossible to distinguish 520 from 530 per minute. In equating proportional rather than absolute (or interval) changes, the logarithmic ordinate models the assumption that behavior rates change in ratios rather than in intervals. This assumption is especially evident in the use of quarter-intercept 'celeration lines to summarize trends.

The relevant statistical questions are whether the logarithmic ordinate preserves homogeneity of variance across changes in level, and whether extensions of 'celeration lines, based on logarithmically transformed data accurately predict the future course of behavior.

Koenig's (1972) doctoral thesis addressed just these questions. From published reports and from a computerized "bank" of human response rate information, he gathered 1,186 sets of data. Each set included between 20 and 29 days of data with an average change of at least 10% per week (x1.10 or ÷1.10 'celeration). He found that trend lines drawn through these data sets, by either quarter-intersect or least squares regression methods, bisected 90% of the points in more than 90% of the cases. He found that changes in range between first and fourth quarters were less than 10% in more than 50% of the cases. And 5 to 7 day projections of variance envelopes from the first 10 to 14 days of data encompassed 70% of the points in more than half the cases. These data suggest that the logarithmic transformation is appropriate for human response rate data. Moreover, Koening (1972) found that the quarter-intersect method was as accurate in summarizing the data as the least squares regression technique.

One of the implications of the logarithmic transformation of response rates is that a given absolute increase represents a greater change from a lower rate than from a higher rate. This is a reasonable assumption. A student who progresses from five to ten problems completed per minute seems, intuitively, to have made greater progress than one who moves from 50 to 55 per minute. The first is a 100% (or x2.00) increase, whereas the latter is merely a 10% (or x1.10) improvement.

On the other side of the transformation issue there is the possibility that a power, or log-log chart may fit human rate data better than the semi-logarithmic chart. Snoddy (1926) originally suggested that the typical shape of the learning curve for highly coherent speed skills such as typing, telegraphy, and industrial assembly is the power function, according to which logarithmic increments in amount of practice result in logarithmic gains in speed. Blackburn (1936) came to a similar conclusion about the effects of practice on the rate of solving mental arithmetic problems. Klemmer (1962) reported data showing that reaction times in simple perceptual-motor tasks follow the same function. And, interestingly enough, some researchers in the precision teaching field have suggested that the log-log scale may actually fit classroom data better than the log-interval Standard Behavior Chart (White & Billingsley, Note 8). They have found that it is possible to make somewhat more accurate predictions and decisions about performances in the proficiency-building stage of learning from data plotted on log-log charts.

It may be that the log-log chart's advantage appears only toward the highest levels of proficiency, where physiological limits, or limits imposed by component skill deficits, create an asymptote. At that point the greater curve-straightening ability of the log-log scale erases such a natural ceiling effect.

For the behavior analyst interested primarily in pragmatics, the log-interval scale has decided advantages. Whereas on the interval calendar base it makes no difference where the course of behavior change begins, on a ratio scale it makes a great deal of difference. In practice, then, log-log charts could not be calendar-synchronized across behavers and behaviors. Although the power function on the log-log scale allows for a simple mathematical model, the power chart lacks many of the useful ratio properties of the Standard Behavior Chart. If at some point it becomes obvious that the power chart is a better fit, then perhaps it will be possible to resolve the practical difficulties. In the mean time, collecting response rate data and charting them on a mechanically powerful chart is a giant step ahead in the study of associative skills. Besides, it should be possible to discover many functional relations, even if the fit is not optimal at all points along the course of skill development.

Assessment, Instruction, and Curriculum

Precision teaching has been largely an oral tradition with relatively few publications in academic journals or books. Most people in the field are teachers or teacher-trainers, of whom few have any motivation to publish. There have been a large number of unpublished manuscripts, theses, and working papers, but even the most recent of them do not reflect the state of knowledge in the field. Lindsley trained a numbers of Ph. D.'s at the University of Kansas, and Pennypacker has done the same at the University of Florida. White, Liberty, Haring, and their associates carry on an active program of research at the University of Washington. But, for the most part, fieldworkers have been the main contributors. School districts in Washington, Oregon, Kansas, Florida, Montana, and Ontario are among the prominent centers of precision teaching activity. What follows is an attempt to outline the major methods and findings of these various groups of people who communicate largely by mouth and with data on the Standard Behavior Chart. Being strictly functionalist in their orientation, they generally express their findings and results as empirical generalizations. Nevertheless, teachers and habilitators of any theoretical persuasion can and do use the Standard Behavior Chart.

Experimental Method

Precision teachers generally use some form of plan sheet (Kunzelmann, Cohen, Hulten, Martin, & Mingo, 1970; White and Haring, 1976) which describes the sequence of antecedents, movements, arrangements, and subsequent events that appear in a particular training procedure. Data-based decisions about individual progress lead to changes in the plan. Thus, precision teaching incorporates the basic ingredients of experimental method. The plans and charts are often organized into a standard file system for a given classroom or child (Binder, Note 9).

Channels Terminology

Many precision teachers use a terminology of channels to categorize tasks. They classify input channels into such categories as "hear," "see," "touch," "mark," "do," "think," and so forth. For example, oral reading is a see/say task, as is naming objects, numbers, or colors. Silent reading is see/think. Spontaneous speech is think/say, whereas answering questions might by see/say or hear/say, depending on whether the questions were written or spoken. These categories are very useful for diagnostic analysis of skills since specific deficits may appear in either input, output, or both. They are also useful in the development of skill taxonomies. (e.g., Kovaks, Note 10).

Basic Measurement Techniques

Precision teaching generally involves daily measurement. Because children as young as five years old can learn to count, time, and chart their own performances (Bates & Bates, 1971), most non-handicapped children measure their own behavior in cooperation with peers and under the supervision of teachers. Most skill assessments are probes ranging from 15 seconds to two minutes in duration, depending on the skill. Usually there are instructional and practice periods in addition to the measurement times, although many skills show satisfactory growth with only a few minutes of practice per day. Assessment periods are fixed durations from day to day, whenever possible. This practice ensures a more reliable assessment.

As much as possible, teaching and measurement procedures are designed so that they do not impose ceilings on performance. For example, spelling tests are given in a "free dictation" mode (Haughton, 1972). Rather than reading a word and then waiting until all children have completed their attempts to spell it, the teacher simply reads the words at a normal speaking frequency and allows children to select words, write them down, and then listen for others. In this way, all students have enough to keep them busy, and they can each demonstrate their own, self-paced proficiency. Many skills are see/write tasks. Worksheets allow students to move at their own pace and to work in what sequence they choose. In this way the precision teaching classroom combines the best aspects of so-called "free" schools and the rigor of data-based instruction. The teacher becomes more of an adivsor to individuals, involving them in the day-to-day decision-making process, than a group lecturer. Students learn self-management, and are motivated by their own progress. Timers and counters of all kinds abound in such classrooms.

Early Discoveries

Relative unimportance of consequences. Since Lindsley and many of his early associates were operant conditioners, they originally emphasized consequences are necessary but not sufficient conditions for skill development (Haughton, 1972). Although precise use of functional consequences can produce acquisition of skill, the early workers found that students would often continue to exhibit rather low rates of performance in spite of fixed ratios of reinforcement and other arrangements. They came to much the same conclusion as Underwood (1966) who said, in summing up a symposium on skill acquisition:

> No participant in the conference seemed impelled to defend the proposition that reinforcement in the classical sense should enter as a major theoretical notion into theories of skill learning. The anal-

ogy between the animal learning laboratory and the skills laboratory can sometimes become quite close, as for example, in the use of information feedback in simple positioning movements. Nevertheless, the consensus seems to be that such feedback is best conceived as a means of defining to the subject the nature of the response to be learned rather than as a reinforcing event in the classical sense.

At least with nonhandicapped children, consequences seemed to be primarily a way of maintaining practice at a task. Some simple methods of differential reinforcement of high rates boosted children's proficiencies, but not nearly so much as careful sequencing of antecedents and the development of prerequisite movement cycles.

<u>Response-response relationships</u>. Haughton (1972) recounts how early precision teachers first became aware of the importance of high levels of proficiency in <u>tool skills</u>, basic movement cycles which constitute lowest common denominators across academic skills. He and his associates understood the basic principles of task analysis, and thus had begun to carefully sequence prerequisite skills in order to maximize positive transfer. But they had not yet learned the importance of high levels of proficiency in the basics. It became clear, for example, that students should be able to think/write <u>1</u>'s and <u>0</u>'s very proficiently in order to be able to progress smoothly through a printing curriculum. They began to see that reading words quickly -- <u>any</u> words -- was an important prerequisite for moving on to new material. Proficient speaking, think/say sounds, is prerequisite to proficient reading. And students should be able to see/say numbers and think/write numbers quickly in order to progress smoothly through an arithmetic sequence. Thus, even though they had been measuring response rates from the very beginning, the early precision teachers took some time to fully appreciate the importance of high rates of responding.

Even though they were measuring rates, they were still thinking in terms of absence/presence of behavior, or percentages. But in the early 1970's they became aware of what have more recently been termed deficit-imposed ceilings (Binder, Note 11).

There are at least four finds of ceilings which impose limits on the growth of proficiency. It is necessary to remove the measurement-imposed ceiling in order to inspect a student's skill proficiency beyond mere acquisition. Changes in procedure remove teacher-imposed ceilings,allowing students to perform at their own pace. It then became obvious that there are deficit-imposed ceilings, created by lack of proficiency in important prerequisite movement cycyles. For example, students who cannot proficiently fill a container with objects will experience difficulty in sorting the objects into different containers. And filling containers may be constrained by deficits in reaching, grasping, placing, or releasing objects. Similarly, low rates in writing or printing letters will constrain the development of proficiency in free-dictation spelling or any other writing task. With physically or mentally handicapped children, it may be that certain irremediable limits, called student-defined ceilings (Binder, Note 11), will impose permanent limits on the growth of certain skills. In that event,it will be necessary to provide prosthetic training, or prosthetic environments (Lindsley, 1964).

In short, just as perceptual-motor skill researchers (Bilodeau & Bilodeau, 1954; Gagne, Baker & Foster, 1950) had found that improving the performance of a less proficient component, even at the expense of practice on a more proficient one, could boost over-all task proficiency, so the precision teachers discovered the importance of proficiency in response-response relationships. This is also consistent with Barrett's (Note 1) finding that some severely retarded subjects could learn discriminations as thoroughly as any normal adult, but that they were deficient in response rate and thus took longer to acquire the discriminations.

<u>Fluency Aims</u>

Precision teachers formalized their findings about the need for high levels of component skill proficiency in terms of <u>fluency aims</u>, or response rate criteria (Haughton, 1972). They began systematically to determine the rates at which students must perform component skills to progress smoothly from acquisition to proficiency, or fluency, in subsequently learned skills. Often charts would indicate that students were having difficulty in developing fluency in particular skills. Stepping back to prerequisite skills and boosting proficiency often eliminated this difficulty on return to the subsequent step in the sequence. Thus, teachers began to discover minimum fluency aims for each step in well-ordered skill sequences. They then were able to develop assessment strategies based on these aims.

Starlin (1972), for example, developed assessment inventories for oral reading, basic computational skills, and spelling. Her strategy was based on three levels of assessment. An <u>inventory level</u> would consist of a relatively large slice of the material, for example, addition sums from 0-9. The <u>screening level</u> included smaller slices within a given inventory, for example, sums from 0 to 5. Finally, the <u>item level</u> included cumulative sets of specific items within a screening, for example, sums involving only the addition of 1, 2, and 3. In reading and spelling, the levels would be determined by phonetic regularities, length of words, and so forth. Starlin learned that certain ranges of performance could guide the teacher in selecting the appropriate level of instruction. In oral reading, she found that a performance of fewer than 20 words per minute indicates the need to <u>slice back</u> to a smaller amount of material, perhaps from the inventory level to a screening level. And a performance of 50 to 100 words per minute indicated that students could practice independently and continue to improve.

<u>Minimum aims for application or transfer</u>. Thus, teachers discovered the need for establishing minimum rate aims to ensure smooth application of component skills to subsequent composite skills. If we think of response rate as a measure of probability, this finding simply elucidates the requirement that skills occur with high probability under appropriate conditions in order for combinations of those skills to occur with high probability. This is perhaps the most basic finding of precision teaching. And it is a principle with broad generality, applicable to a range of issues in the study of associative skills. It is founded on the use of response rate as a measure of such skills.

<u>Endurance and retention</u>. High levels of fluency also appear to be necessary for endurance, or maintenance of fluency over comparatively long periods of performance (Haughton, Note 12). After students develop high levels of proficiency in short assessment periods, they become more capable of practicing on their own for longer periods of time. Haughton has estimated that, to be capable of practicing skills without supervision, students should be able to perform those skills at at least one half the fluency standard. Otherwise, they will decrease their performance rates in the absence of supervision. Endurance and the maintenance of "attention" to task appear closely related, as suggested in Chapter II. Haughton (Note 12) claims that retention is also related to fluency, and White and Haring (1976) suggest that failure to maintain skills may be due to inadequate levels of proficiency. There may be two dimensions to these related observations. First, high probabilities of performance may be related to retention in the traditional sense. That is, skills which are well established in the first place are less likely to be for-

gotten during a period of non-use.

High levels of fluency probably make their greatest contribution to retention by enabling students to maintain skills through application in higher skills. Barrett (1977) suggested that in a well-designed skill sequence, "since every skill in the sequence is a component of or a contributor to the skill above it, skills learned early in the sequence are, presumably, being used repeatedly and therefore should not disappear from the behaver's repertoire."

Norms, fluency, and mastery. Certain levels of performance are necessary for efficient, effective use in appropriate applications. People who read accurately, but slowly, experience difficulty when they have to use reading skills in educational or occupational settings. This is a particularly good reason for moving beyond mere accuracy measurement and for removing teacher-imposed ceilings. Furthermore, slow readers do not generally read for recreation and therefore do not maintain whatever skill they once attained. Similarly, retarded persons who are capable of making change accurately, but very slowly, are unlikely to maintain the skill. Instead, friendly or impatient shopkeepers will probably do it for them. A definition of _fluency_, then, is the rate at which skills are effective and relatively efficient. _Dysfluency_ is defined as a rate at which skills are ineffective.

One way of determining minimum fluency standards is to assess people assumed to be normal in their performance of skills. Thus, precision teachers have gathered data to establish _norms_ (Barrett, 1979). Adult norms are the usual standards of fluency, although teachers often use _peer norms_ as ways of establishing age-appropriate aims (Child Demonstration Programs, Note 2; White & Haring, 1976).

Because norms seldom represent levels of skill _mastery_, another approach (Haughton, Note 12) seeks to discover the levels at which truly exceptional individuals perform skills. With these as aims, it may be possible to develop true mastery in large numbers of individuals. In fact, whereas Haughton (1972) previously set 100 words per minute with fewer than 2 errors as an aim for beginning readers, he now suggests an aim of 200 to 300 per minute as an appropriate level (Haughton, Note 12). Although it may be dangerous to set aims too high, it is certainly dangerous to set aims too low.

Decision Rules

The Standard Behavior Chart forms a basis for educational decision-making. But it does not make the decisions. Teachers must do so. And in order to increase the frequency with which teachers make data-based instructional decisions, a number of precision teachers, especially Kathleen Liberty and Owen White, have been working to develop a series of decision rules based on patterns of data points on the chart.

Freschi's (1974) discussion of precision teaching represents the least formal approach to decision-making. Teachers simply inspect charted data, noting whether rates of correct and incorrect responses are changing in the desired directions, or not. If they are, the teacher continues the current plan. If not, the plan is changed. Although daily data collection and data-based decisions in this form represent a major advance over typical educational procedures, many precision teachers nonetheless make very few decisions, allowing procedures to continue without change even when ineffective.

Noting this low frequency of decision-making, Liberty (Note 13) invented dynamic aims, or progress aims. This procedure involves specification of a 'celeration, or rate of change, according to which teachers can judge whether adequate progress is occurring. Liberty analyzed data on several hundred charts including all kinds of skills and children of all ages. She found that of those students whose programs produced some progress, 50% showed growth of x1.33 or better on acceleration targets and ÷1.46 for deceleration targets. About 53% progressed at x1.25 for acceleration, 66% at ÷1.25 for deceleration. She suggested 1.25 as a minimum 'celeration. After collecting three days of data and finding the midpoint, teachers using dynamic aims draw a 'celeration line from that point toward the fluency aim on the chart. If on more than three days in a row, the student's rate falls below (or above, in the case of decelerations) the dynamic aim, the teacher changes the plan. Teachers may learn that some children are able to change faster than 25% per week, and thus establish individualized progress aims for such children. But the x1.25 or ÷1.25 rule is a useful first choice for minimum 'celeration.

The use of dynamic aims substantially increased the frequency of decision-making (Haring & Liberty, Note 14), but it still does not suggest what changes are likely to produce the desired effect. Should the teacher alter the consequences, slice back on the content, step back to an easier skill, or alter the instructional procedure? A five-year research project, currently under way at the University of Washington, is designed to solve this problem (Haring, Note 15). With Liberty as its coordinator, this project has specified and repeatedly revised a set of rules according to which teachers can make the change most likely to succeed. After drawing split-middle 'celeration lines through patterns of correct and error rates on the chart, teachers refer to these rules, which suggest the appropriate change for a given pattern of accelerations and/or decelerations of corrects and errors (Haring, Note 16). Over the several years of the project, the researchers have collected data on the successes and failures of each successive revision of the rules. The most recent set of rules predicted the effects of plan changes in an average of 68% of the cases, across the various categories of change.

Although the University of Washington group has also attempted to derive such rules on the basis of percentage correct data, they have been unable to do so.

Diagnostic/Prescriptive Assessments

Establishment of fluency aims and the development of decision-making rules has allowed precision teachers to carry out precise diagnostic/prescriptive assessments. The simplest form of assessment involves a few days' collection of data, and comparison with fluency aims and/or peer ranges of performance across a selection of skills. Even one-day samples ("snapshots") are useful in this regard (e.g., Haughton, Note 17). It is possible to apply this form of assessment within a classroom, in an entire school, or across several school districts (Child Service Demonstration Programs, Note 2).

However, it is far more useful to select a set of component, prerequisite, and composite skills in a given curriculum area, and then assess students for 5 to 10 days on those skills. In that way, not only levels of proficiency, but also 'celerations over the duration of the assessment period, allow teachers to decide where to place students in a given curriculum and what deficits to remediate (White & Haring, 1976). This approach represents a substantial improvement over traditional forms of standardized assessment. It allows students to demonstrate proficiency levels on specific skills, with repeated opportunities, over a reasonable period of time. And it compares their actual performances to objective norms of performance on those skills rather than to very general norms, established according to psychometric principles.

Instructional Implications of Rate Measurement: An Example

In most elementary schools reading is either implicitly or explicitly assessed on a percentage correct scale. Accordingly, students first learn to read a story or list of words, and then when they have exhibited the ability to do so, pass on to the next story or set of words. In fact, since most educational methods do not even involve criterion-referenced assessment, students pass on to the next story whether or not they were able to accurately read the previous one. But in any case, teachers do not generally advance students on the basis of fluency aims. As a result, first-graders may pass through a list of books, still reading at between 30 and 80 words per minute. Doehring (1976), for example, found that average first graders read about 33 <u>syllables</u> per minute of random words on worksheets, and about 46 syllables per minute in text. These relatively low rates of reading may create severe problems in retention and comprehension.

A precision teacher would require students to reach relatively high fluency aims on their first selection, between 200 and 300 <u>words</u> per minute (Menninga and Brough, Note 18). Having reached this rate on a given passage, the student may be <u>reciting</u> it from memory. But the student is nonetheless reading it, responding with appropriate movement cyles to textual stimuli. And as successive passages increase vocabulary, it will become impossible to memorize them.

With each change to a new passage, there will be a rate decrement and an increase in errors. But relatively quickly, the performance will return to fluency. Thus, from the very beginning, a child's oral reading will be maintained at near adult levels of fluency. Precision teaching emphasizes fluency as a prerequisite for increasing content. The traditional approach emphasizes content, often at the expense of fluency.

<u>Precision Teaching with the Retarded</u>

Although it was work with the retarded that originally inspired the development of the Standard Behavior Chart (Lindsley, 1964), and many of the first precision teachers worked with the retarded (Galloway, 1972), most of the important advances in precision teaching until recently have occurred in classrooms for regular and mildly handicapped students (White & Haring, 1976). There are several reasons for this.

First, severely retarded students generally have not learned to chart their own behavior. Using the Standard Behavior Chart in classrooms for such students therefore involves significantly more work for teachers than does work with students who do their own timing, counting, and charting.

Second, it is more difficult to remove teacher-imposed ceilings from instructional procedures for the severely retarded. Whereas regular children can work with worksheets and other self-presented materials from the very beginning, painstaking procedures are often needed to move severely retarded students toward self-paced procedures (Pease and George, Note 19). Therefore, teachers are not readily convinced that proficiency is possible for such students.

Third, severely retarded and otherwise handicapped students often exhibit substantial day-to-day variability in performance (Haring, Note 6). This may result from various neurological impairments, physical illness, or the failure of teachers to obtain instructional compliance.

Finally, only recently have precision teachers begun to look carefully at critical tool skills for many of the fine and gross motor performances that constitute the curricula in most classrooms for the severely retarded. Haughton (Note 12) has observed that deficits in reaching, pointing, touching, grasping, placing and releasing impose ceilings on most performances of these students. He has suggested that teachers return to these basic motor components, along with basic body control movements, in their programming for the severely retarded. Haughton and his associates (Kovacs, Note 10) have begun to establish fluency aims for these skills.

Precision Teaching as Functional Analysis

Above all, precision teachers are functionalists. From its conception (Lindsley, 1972) precision teaching has been measurement-based and atheoretical, following the tradition of functional analysis established by Skinner (1950). The establishment of fluency aims, curriculum sequences, and decision rules have developed directly from empirical data, with no intervening levels of hypotheses or constructs. In this sense, precision teaching differs from all other approaches to education. In fact, Lindsley (1971) has denied that it is an "approach." Rather, he says, it is a precise way of monitoring the effects of whatever procedures a teacher might choose. It is simply the daily measurement of response rates in the classroom. In this sense, it also differs from a great deal of the theory-based research that educators and psychologists have conducted on various aspects of learning and skill acquisition.

LaBerge and Samuels (1974), for example, present a theory of "automatic information processing in reading." The following quotation contrasts sharply with the strictly functional analysis of precision teaching:

> In the present model it is assumed that all well-learned stimuli are processed upon presentation into an internal representation, or code, regardless of where attention is directed at the time . . . . It is assumed that we can only attend to one thing at a time, so long as no more than one requires attention.... Our criterion for deciding when a skill or sub-skill is automatic is that it can complete its processing while attention is directed elsewhere. (p. 294-295)

In various flowcharts, the authors construct an elaborate system of coders, detectors, and processors. Like Fitts (1964) they use a computer model of "programs and sub-routines" to account for observed functional relationships.

They come to the same conclusions as precision teachers as to the difference between accuracy and proficiency at different stages in the learning process, and they use time as a measurement dimension. "Further training beyond the accuracy criterion must be provided", they say, "if the association is to occur without attention.... Latency serves as the critical indicator of automatic associative processing" (p. 307). In this type of theorizing, hypothetical processes appear more important than data. Thus they speak of measurement as <u>indicating</u> an underlying state of automaticity, rather than as simply one element in a functional relationship.

Like precision teachers, LaBerge & Samuels (1974) find that practice beyond accuracy is important. And they remark that:

> The finding that unfamiliar letters improved with practice more than did familiar letters offers support for the hypothesis that something is being learned about the unfamiliar letters over the days of training.... It is evident that the latency difference of a naming response to the new and old letters is quite large at first and converges over days of training....when accuracy appears to be stationary (p. 303).

The finding that learning continues beyond mere accuracy led them to "The main conclusion....that what is being improved with practice is automaticity" (p. 304). But what <u>is</u> "automaticity"? As Skinner has frequently pointed out (1950, 1974, 1978), theoreticians often move the environment inside the organism, accounting for an observation - - in this case a short latency response which is relatively immune to distraction - - in terms of an internalized synonym for the behavior. In this context, we might ask whether such move creates no obvious practical predictive advantage.

LaBerge & Samuels (1974) agree with precision teachers on the importance of feedback:

> Since the important growth of automaticity takes place after the subject has achieved accuracy, overt feedback for correct and incorrect

responses may be redundant because at this stage of learning the subject knows when he is correct or not. However, there is another type of feedback which may affect the rate of automaticity learning. While learning proceeds toward the automatic level, it might be appropriate to inform the subject of the time it took to execute his response. (p. 317)

This, of course, is exactly the kind of feedback provided to students who count, time, and chart their own behavior. LaBerge and Samuels go on to say:

> Of course, latency feedback for a response to a particular stimulus will not be meaningful by itself, but must be related to some criterion baseline. For example, the time it takes to identify a new word should be compared to the time it takes to identify a word that is already at the automatic level. (p. 317)

This, of course, if the function of fluency aims in precision teaching.

Relating proficiency to skill-sequencing, LaBerge and Samuels note that "the fluent reader has presumably mastered each of the subskills at the automatic level" (p. 318). Continuing in this vein, they say:

> We are tempted to generalize to classroom routines in elementary schools in which letter naming is directly taught and tested only up to the accuracy level. A child may be quite accurate in naming or sounding the letters of the alphabet, but we may not know how much attention it costs him to do it. This kind of information could be helpful in predicting how easily he can manage new learning skills which build on associations he has already learned.... One should take into account the amount of attention required by these subskills as a part of the readiness criterion.

We might ask, why not simply examine how fluently a student must be able to perform a subskill to progress smoothly through the next curriculum level? Does the concept of "attention" add anything here?

In discussing "levels of processing," LaBerge and Samuels invoke the notion of chunking to account for increased fluency. They observe that "to encourage chunking, we may have to relax the demand for accuracy" (p. 319). This speculation is strikingly consistent with Haughton's (1972) finding that until students reach a reading rate of about 50 words per minute, attempts to decelerate errors are largely ineffective. After that point, it is relatively easy to reduce error rates.

Finally, LaBerge and Samuels (1974) remark:

> Throughout this paper, we have stressed the importance of automaticity in performance of fluent reading. Now we turn to a consideration of ways to train reading subskills to automatic levels. Unfortunately, very little systematic research has been directed specifically to this advanced stage of learning. (p. 314)

Doehring (1976), writing in a similar vein, asserts that "a deeper understanding of the learning process is necessary before reading instruction can be improved" (p.42). He says that among the greatest difficulties are "the problems of describing the accuracy and speed of responses in the same quantitative terms." (p. 42)

As this review of precision teaching indicates, there has been a good deal of work on the development of fluency in reading, although little has been published. The work has focused, not on the "underlying processes," but on the discovery of functional relationships. It has used measures of response rates and accuracy ratios, which solve the problems of simultaneously quantifying accuracy and speed and which are also more convenient and less costly than latency measures.

In the continuous feedback loop between basic and applied research, it might serve the interests of both researchers and educators to emulate the functional approach of precision teachers. Their work has uncovered a number of important phenomena that deserve closer inspection from basic researchers.

On Competency Testing

There is currently great concern about the failure of schools to teach students reading, writing, and computational skills. Many have suggested that all students should take competency tests at the end of high school - - a proposition that might only result in proof of failure long after it is possible to do anything about it. Precision teaching, incorporating daily, data-based decisions about instructional effectiveness, would seem a more appropriate antidote. The alternative seems to be continued wave after wave of theoretical explanations about how learning occurs, and why it doesn't. Without an empirical approach, based on daily measurement, it seems likely that functional ill-iteracy, now estimated to afflict at least 20% of the adult population in this country, will continue to increase.

Summary and Conclusions

Precision teaching is based on functional analysis of behavior, originally developed in the free operant conditioning laboratory. It involves the measurement and charting of response rates on the Standard Behavior Chart, a powerful communication device.

The Standard Behavior Chart is a six-cycle semilogarithmic graph. Its ordinate is logarithmic while its 140-day calendar base is linear. It is possible to plot durations and latencies on the chart as reciprocals, enabling the charter to inspect relations between fluency and session length, among other things. The ordinate spans a range of response rates from one per day to 1000 per minute. Its logarithmic scale makes it possible to examine accuracy, variability, and trend independent of rate. It also enables the charter to summarize data in terms of a number of ratios or multipliers.

Precision teaching is primarily a measurement system. On the basis of that system, teachers have discovered a variety of empirical regularities. In particular, they have discovered the importance of developing high levels of fluency in component skills. Fluency is related to transfer, retention, or maintenance, and to endurance of skills. These findings lead to the establishment of fluency aims and decision-making rules and to diagnostic/prescriptive skill analysis.

In contrast to an information processing approach, precision teaching has achieved a great deal in classrooms, without ornate theoretical structures. As a basic methodology, precision teaching seems eminently applicable to laboratory study of associative skills.

References

Abrams, S. G., & Zuber, B. L. Some Characteristics of information processing during reading. Reading Research Quarterly, 1972, 8(1), 40-51.

Abrams, Jack A. Motor behavior. In Melvin Marx (Ed.), Learning Processes. New York: The Macmillan Co, Inc., 1972.

Atkinson, R.C., and Shiffrin, R.M. Human memory: A proposed system and its control processes. In K.W. Spence & J.T. Spence (Eds.), The Psychology of Learning and Motivation. New York: Academic Press, 1968.

Barrett, B. H. Behavior analysis. In J. Wortis (Ed.), Mental Retardation and Developmental Disabilities Vol. 9. New York: Brunner/Mazel, 1977.

Barrett, Beatrice H., & Lindsley, Ogden R. Deficits in acquisition of operant discrimination and differentiation shown by institutionalized retarded children. American Journal of Mental Deficiency, 1962, 67(3), 424_436.

Barrett, Beatrice H. Acquisition of operant differentiation and discrimination in institutionalized retarded children. American Journal of Orthopsychiatry, 1965, 35(5), 862-885.

Barrett, Beatrice H. Communitization and measured message of normal behavior. In R. York & E. Edgar (Eds.), Teaching the Severely Handicapped Vol. 4. Columbus: Special Press, in press.

Bates, S., & Bates, D. F. "… and a child shall lead them": Stephanie's chart story. Teaching Exceptional Children, 1971, 3, 111-113.

Beck , C. H. M. Paced and self-paced serial simple reaction time. Canadian Journal of Psychology, 1963, 17(1), 90-97.

Bilodeau, Edward A., & Bilodeau, Ina McD. The contribution of two component activities to the total psychomotor task. Journal of Experimental Psychology, 1954, 47(1), 37-46.

Blackburn, J. M. Acquisition of Skill: An analysis of learning curves. Great Britain, I.H.R.B. Report No. 73, 1936.

Braam, Leonards S., & Berger, Allen. Effectiveness of four methods of increasing reading rate, comprehension and flexibility. Journal of reading, 1968, 11(5), 346-352.

Briggs, George E. Transfer of training. In Bilodeau, E. A., & Bilodeau, Ina McD. (Eds.), Principles of Skill Acquisition. New York: Academic Press, 1969.

Butler, D.C., & Peterson, D.E. Learning during "extinction" with paired associates. Journal of Verbal Learning and Verbal Behavior, 1965, 4, 103-106.

Cattell, J. McKeen. The time it takes to see and name objects. Mind, 1886, 11, 63-65.

Connolly, Kevin, Brown, Kathleen, & Bassett, Eryl. Developmental changes in some components of a motor skill. British Journal of Psychology, 1968, 59(3), 305-314.

Cofer, Charles N. Verbal Learning. In Melvin Marx (Ed.), Learning Processes. Mew York: The Macmillan Co., Inc., 1972.

Demckla, Martha Bridge, & Rudel, Rita. Rapid "automatized" naming of pictures, objects, colors, letters, and numbers by normal children. Cortex, 1974, 10(2), 186-202.

Doehring, Donald G. Acquisition of rapid reading responses. Monographs of the Society for research in Child Development, 1976, 41(2, Serial No. 165).

Dresslar, F. B. Some Influences which affect the rapidity of voluntary movements. American Journal of Psychology, 1892, 4, 514-527.

Duncan, Ann Dell. Behavior rates of gifted and regular elementary school children. (Doctoral dissertation University of Kansas, 1969). Dissertation Abstracts International, 1969, 29(10-A), 3520.

Eaton, Ira E. The relationship between perceptual-motor ability and reading success. (Doctoral dissertation, St. Louis University, 1975). Dissertation abstracts, 1975, 6, 3562-A.

Ebbinghaus, H. Memory: A Contribution to Experimental Psychology. (H. A. Ruger & C.E. Bussemius, trans.). New York: Teachers College, Columbia University, 1885.

Edwards, D.D., & Edwards, J. S. Fetal Movement: development and time course. Science, 1970, 169, 95-97.

Eimas, P. D., & Zeaman, D. Response speed change in an Estes' paired associate "miniature" experiment. Journal of Verbal Learning and Verbal Behavior, 1963, 1, 384-388.

Ellis, Henry C. Transfer and Retention. In Melvin Marx (ed.), Learning Processes. New York: The Macmillan Co., Inc., 1972.

Estes, W. K., & Skinner, B. F. Some quantitative properties of anxiety. Journal of Experimental Psychology, 1941, 29, 390-400.

Eysenck, S. B. G. Retention of a well-developed motor skill after one year. Journal of Genetic Psychology, 1960, 63, 267-273.

Ferster, C. B., & Skinner, B. F. Schedules of Reinforcement. New York: Appleton-Century Crofts, 1957.

Fitts, Paul M. Perceptual-motor skill learning. In Arthur W. Melton (Ed.), Categories of Human Learning. New York: Academic Press, 1964.

Fitts, P. M., & Peterson, J. R. The Information capacity of discrete motor responses. Journal of Experimental Psychology, 1964.

Fitts, Paul M. The information capacity of the human motor system in controlling the amplitude of movement. Journal of Experimental Psychology, 1954, 47(6), 381-391.

Fleishman, Edwin A. Comments on Professor Jones' paper. In Bilodeau, Edward A. (Ed.), Acquisition of Skill. New York: Academic Press, 1966.

Freschi, D. F. Where we are. Where we are going. How we're getting there. Teaching Exceptional Children, 1974, 6, 89-97.

Gagne, Robert M. The Conditions of Learning. New York: Holt, Rinehart & Winston, Inc., 1970.

Gagne, R. M., & Foster, Harriet. Transfer of Training from practice on components in a motor skill. Journal of Experimental Psychology, 1949, 39, 47-68.

Gagne, R. M., Baker, K. E., & Foster, H. Transfer of discrimination training to a motor task. Journal of the Experimental Psychology, 1950, 40, 314-328.

Gagne R. M. Task analysis:-- its relation to content analysis. Journal of Educational Psychology, 1974, 11, 11-18.

Gagne, R. M., & Briggs, L. J. Principles of Instructional Design. New York: Holt, Rinehart & Winston, 1974.

Galloway, Charles. Precision parents and the development of retarded behavior. In June B. Jordan, & Lynn S. Robbins (Eds.), Let's Try Doing Something Else Kind of Thing. Arlington, Virginia: Council for Exceptional Children, 1972.

Goldiamond, Israel. Machine definition of ongoing silent and oral reading rate. Journal of the Experimental Analysis of Behavior, 1962, 5(3), 363-367.

Goss, A.E. Manifest strengthening of correct responses of paired-associate under post criterion zero percent occurrence of response members. Journal of Genetic Psychology, 1965, 72, 135-144.

Hall, John F. Verbal Learning and Retention. Philadelphia: J.B. Lippincott Co., 1971.

Hall, R. F., & Wenderoth, P. M. Effects of number of responses and recall strategies on parameter values of a paired-associate learning model. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 29-37.

Haughton, Eric. Aims: Growing and Sharing, In June B. Jordan, & Lynn S. Robbins (Eds.), Let's Try Doing Something Else Kind of Thing. Arlington, Virginia: Council for Exceptional Children, 1972.

Heinz, Robert D., & Eckerman, David A. Latency and frequency of responding under discrete-trial fixed-interval schedules of reinforcement. Journal of the Experimental Analysis of Behavior, 1974, 21(2), 341-355.

Herrnstein, Richard J., & Boring, Edwin G. (Eds.). A Source Book in the History of Psychology. Cambridge: Harvard University Press, 1965.

Hull, C. L. Principles of Behavior. New York: Appleton-Century, 1943.

Hull, C. L. The place of innate individual and species differences in a natural-science theory of behavior. Psychological Review, 1945, 52, 55-60.

Horton, David L., & Kjeldergaard, Paul M. An experimental analysis of associative factors in mediated generalization. Psychological Monographs: General and Applied, 1961, 75(11, Whole No. 515).

Irion, Arthur L. A Brief history of research on the acquisition of skill. In Edwards A. Bilodeau (Ed.), Acquisition of Skill. New York: Academic Press, 1966.

Irion, Arthur L. Historical introduction. In E. A. Bilodeau & Ina McD. Bilodeau (Eds.), Principles of Skill Acquisitions. New York: Academic Press, 1969.

Jenkins, James J. Mediated Associations: Paradigms and situations. In Charles N. Cofer & Barbara S. Musgrave (Eds.), Verbal Behavior and Learning. New York: McGraw-Hill Book Co., Inc., 1963.

Johnston, J. M., & Pennypacker, H. S. A behavioral approach to college teaching. American Psychologist, 1971, 26, 219-244.

Judd, Wilson A., & Glaser, Robert. Psychonomic Bulletin, 1967, 1(2), 11.

Judd, Wilson A., & Glaser, Robert. Response latency as a function of training method, information level, acquisition, and overlearning. Journal of Educational Psychology, 1969, 60(4, Pt. 2), 30.

Keller, F. S. The Phantom plateau. Journal of the Experimental Analysis of Behavior, 1958, 1(1), 1-13.

Keller, L., Thomson, W. J., & Tweedy, J. R. Efforts of reinforcement intervals in pair-associate learning. Journal of Experimental Psychology, 1967, 73(2), 268-277.

Klemmer, E. T. Communication and human performance. Human Factors Journal, 1962, 4, 75-79.

Kunzelmann, H. P., Cohen, M., Hulten, W., Martin, G., & Mingo, A. Precision Teaching: An Initial Training Sequence. Seattle: Special Child Publications, 1970.

LaBerge, David, & Samuels, S. Jay. Toward a theory of chronic psychotics as revealed by free operant conditioning methods. Diseases of the Nervous System Monograph, 1960, 21, 66-78.

Lenneberg, Eric H. Speech as a motor skill with special reference to nonaphasic disorders. In Ursula Bellugi & Roger Brown (Eds.), The acquisition of language. Monographs of the Society for Research in Child Development, 1964, 29(1, Serial No. 92).

Leonard, J. A., & Conrad, R. Maintenance of high accuracy without augmented feedback. Nature, 1963, 199, 512-513.

Lindsley, Ogden R. Direct measurement and prosthesis of retarded behavior. Journal of Education, 1964, 147, 62-81.

Lindsley, Ogden R. From Skinner to precision teaching: The child knows best. In June B. Jordan, & Lynn S. Robbins (Eds.), Let's Try Doing Something Else Kind of Thing. Arlington, Virginia: Council on Exceptional Children, 1972.

Maliphant, R., Supramaniam, S., & Saraga, E. Acquiring skill in reading: a review of experimental research. Journal of Child Psychology and Psychiatry, 1974, 15(3), 175-185.

Marx, Melvin. Learning Processes. New York: The Macmillan Co., Inc., 1972.

McGeoch, J.A. The acquisition of skill. Psychological Bulletin, 1927, 24, 437-466.

Milward, R. Latency in a modified paired-associate learning experiment. Journal of Verbal Learning and Verbal Behavior, 1964, 3, 309-316.

Morton, John The effects of context upon speed of reading, eye movements and eye-voice span. Quarterly Journal of Experimental Psychology, 1964, 16(4), 340-354.

Naylor, James C., & Briggs, George E. Long-term Retention of Learned Skills: A Review of the Literature. United States Air Force ASD Technical Report (No. 61-390), 1961.

Osgood, Charles E. Meaningful similarity and interference in learning. Journal of Experimental Psychology, 1946, 36(4), 277-301.

Otto, W., & Fredericks, R.C. Relationship of reactive inhibition and reading attainment. Journal of Educational Psychology. 1963, 54, 227-230.

Pennypacker, H. S., Koenig, C. H., & Lindsley, O. R. Handbook of the Standard Behavior Chart. Kansas City, Kansas: Precision Media, 1972.

Peterson, Lloyd R. Paired-associate latencies after the last error. Psychonomic Science, 1965, 2, 167-168

Postman, Leo. Transfer, interference and forgetting. In Kling, J. W., & Riggs, Lorrin A. (Eds.), Experimental Psychology. New York: Holt, Rinehart & Winston, Inc., 1971.

Richardson, J., & Gropper, Mitsi S. Learning during recall trials. Psychological Reports, 1964, 15, 551-560.

Roberts, M., Bondy, A., Mira, M., & Cairns, G. Continuous tracking of behavioral tracking in infants. Journal of Genetic Psychology, 1978, 132(2d Half), 225-60.

Runquist, W. N. Verbal behavior. In J. B. Sidowski (Ed.), Experimental Methods and Instrumentation in Psychology. New York: McGraw-Hill, 1966.

Shackel, B. Pilot study in electro-oculography. British Journal of Ophthalmology, 44, 89-113.

Schmidt, Bernard. Changing patterns of eye movement. Journal of Reading, 1966, 9(6), 379-385.

Skinner, B. F. The Behavior of Organisms. New York: Appleton-Century Crofts, 1938.

Skinner, B. F. Science and Human Behavior. New York: The MacMillan Co., 1953.

Skinner, B. F. The Technology of Teaching.

Skinner, B. F. Walden Two. New York: Macmillan, 1948.

Skinner, B. F. Are theories of learning necessary? Psychological Review, 1950, 57, 193-216.

Skinner, B. F. The analysis of behavior. American Psychologist, 1953c, 8, 69-79.

Skinner, B. F. A case history in scientific method. American Psychology, 1956, 11, 221-223

Skinner, B. F. About Behaviorism. New York: Knopf, 1974.

Skinner, B. F. Why I am not a cognitive Psychologist. In B. F. Skinner, Reflections on Behaviorism and Society. New York: Prentice Hall, 1978.

Shoddy, G. S. Learning and stability. Journal of Applied Psychology, 1926, 10, 1-36.

Spring, Carl. Encoding speed and memory span in dyslexic children. Journal of Special Education, 1976, 10, 35-40.

Spring, Carl. Naming speed as a correlate of reading ability and sex. Perceptual and Motor Skills, 1975, 41(1), 134.

Starlin, Ann. Sharing a message about curriculum with my teacher friends. In June B. Jordan, & Lynn S. Robbins (Eds.), Let's Try Doing Something Else Kind of Thing. Arlington, Virginia: Council on Exceptional Children, 1972.

Stennett, R. G., Smythe, P. C., Pinkney, June, & Fairbairn, Ada. The relationship of eye movement measures to psychomotor skills and other elemental skills in learning to read. Journal of Reading Behavior, 1972-73, 15(1), 1-13.

Stevens, Joseph P. Applications of power functions to perceptual-motor learning. Journal of Experimental Psychology, 1964, 68(6), 614-616.

Stevens, Joseph C., & Savin, Harris B. On the form of learning curves. Journal of the Experimental Analysis of Behavior, 1962, 5(1), 15-18.

Taylor, S.E., Frackenpohl, H., & Pattee, J.L. Grade level norms for the components of the fundamental reading skill. Bulletin No. 3, New York: Huntington, Educational Development Laboratories, 1960.

Taylor, Stanford E. Eye movements in reading: facts and fallacies. American Educational Research Journal, 1965, 2(4), 187-202.

Thalberg, Staton P. Reading rate and immediate versus delayed retention. Journal of Educational Psychology, 1967, 58(6, Pt. 1), 373-378.

Theios, John. Reaction time measurements in the study of memory processes: Theory and data. In G. H. Bower (Ed.), The Psychology of Learning and Motivation, Vol. 7. New York: Academic Press, 1973.

Tinker, Miles A. Devices to improve speed of reading. Reading Teacher, 1967, 20(7), 605-609.

Underwood, Benton J. Motor skills learning and verbal learning. In Edward A. Bilodeau (Ed.), Acquisition of Skill. New York: Academic Press, 1966.

Wainer, Howard. Robust statistics. Journal of Educational Statistics, 1976, 1, 258-312.

White, Owen R., & Harring, Norris G. Exceptional Teaching. Columbus: Charles Merrill, 1976.

White, Owen R., & Liberty, Kathleen A. Behavioral assessment and precise educational measurement. In Norris G. Haring, & Richard L. Schielfelbusch (Eds)., Teaching Special Children, 1976.

White, Sheldon H., Spiker, Charles C., & Holton, Ruth. Associative transfer as shown by response speeds in motor paired-associate learning. Child Development, 1960, 31, 609-616.

Wolfe, Dael. Training. In S. S. Stevens (Ed.), Handbook of Experimental Psychology. New York: John Wiley & Sons, 1951.

Wickens, C. D. Temporal Limits of human information processing: a developmental study. Psychological Bulletin, 1974, 81, 739-755.

Williams, Joanna P. Learning to read: A review of theories and models. Reading Research Quarterly, 1973, 8(2), 121-126.

Wright, Logan, & Willis, Carol. Reminiscence in normal and defectives. American Journal of Mental Deficiency, 1969, 73(5), 700-702.

Reference Notes

1. Barrett, Beatrice H. Annual Report 1972-1973. Waltham, Massachusetts: Behavior Department, Walter E. Fernald State School, 1973.
2. Child Service Demonstration Programs Progress Report IV. Tacoma, Washington: Intermediate School District N. 111, 1974.
3. White, O. R. The "split-middle": A "quickie" method of trend estimation. Working Paper No. 1, University of Oregon: regional Center for Handicapped Children, 1971.

4. Koenig, C. H. Charting the Course of Future Behavior. Unpublished doctoral dissertation, University of Kansas, 1972.

5. White, O. R. The prediction of human performances in the single case: An examination of four techniques. Working Paper No. 15, University of Oregon: Regional Center for Handicapped Children, 1972.

6. Haring, Norris G. An investigation of phases of learning and facilitating instructional events for the severely handicapped, Annual Progress Report, University Of Washington, Experimental Education Unit, 1976-77.

7. Haughton, Eric. Personal communication, October 27, 1978.

8. White, Owen, Billingsley, Feliz, & Munson, Robin. Evaluating the daily progress of severely and profoundly handicapped pupils. Paper presented at the meeting of the American Association for the Education of the Severely and Profoundly Handicapped, Baltimore, October, 1978.

9. Binder, C. V. Guide to a Standard File System. Waltham, Massachusetts: Behavior Prosthesis Laboratory, Walter E. Fernald State School, 1979.

10. Kovacs, Mary (Ed.), A Practical Taxonomy of Normal Body Control and Beginning Skills. Ontario: Hastings County Board of Education, 1978.

11. Binder, C. V. Four kinds of Ceilings. In C. V. Binder (Ed.), Data-Sharing (Waltham, Massachusetts: Behavior Prosthesis Laboratory, Walter E. Fernald State School), October, 1978.

12. Haughton, Eric. Untitled contribution. In C. Binder (Chair), More parameters of pupil freedom. Symposium presented at the meeting of the American Association for the Education of the Severely and Profoundly Handicapped, Baltimore, October, 1978.

13. Liberty, K. A. Data Decision rules. Unpublished working paper, Experimental Education Unit, University of Washington, 1975.

14. Haring, Norris, & Liberty, Kathleen. "What do I do when?" Stages of Learning and facilitating instructional events. Paper presented at the meeting of the American Association for the Education of the Severely and Profoundly Handicapped, Baltimore, October, 1978.

15. Haring, Norris G. Field initiated research studies: An investigation of learning and instructional hierarchies in the severely/profoundly handicapped -- Renewal application for the fifth year. University of Washington, Experimental Education Unit, 1979.

16. Haring Norris G. An Investigation of phases of learning and facilitating instructional events for the severely handicapped, Annual Progress Report. University of Washington, Experimental Education Unit, October, 1978.

17. Haughton, Eric (Ed.). Snapshot Notes. Ontario: Hastings County Board of Education, 1977.

18. Menninga, Rita, & Brough, Mary. Screening grade one – three. Ontario: Hastings County Board of Education, June, 1978.

19. Pease, D., & George, F. Flashcards to worksheets: Transitional training in normalization of academic behavior. In Instructional Development and Behavior Analysis Program, Annual Report. Waltham, Massachusetts: Behavior Prosthesis Laboratory, Walter E. Fernald State School, 1975.