Behavioral Fluency: Evolution of a New Paradigm

Carl Binder

Precision Teaching and Management Systems, Inc.

Behavioral fluency is that combination of accuracy plus speed of responding that enables competent individuals to function efficiently and effectively in their natural environments. Evolving from the methodology of free-operant conditioning, the practice of precision teaching set the stage for discovering about relations between behavior frequency and specific outcomes, notably retention and maintenance of performance, endurance or resistance to distraction, and application or transfer of training. The use of frequency aims in instructional programming by Haughton and his associates led to formulation of empirically determined performance frequency ranges that define fluency. Use of fluency-based instructional methods has led to unprecedented gains in educational cost-effectiveness, and has the potential for significantly improving education and training in general. This article traces the development of concepts, procedures, and findings associated with fluency and discusses their implications for instructional design and practice. It invites further controlled research and experimental analyses of phenomena that may be significant in the future evolution of educational technology and in the analysis of complex behavior.

Key words: fluency, behavior frequency, precision teaching, automaticity, instructional design, free operand

Fluency-based education and training programs have produced some of the most dramatic results in the history of behaviorally oriented instruction. During the 1970s, the Precision Teaching Project in Great Falls, Montana (Beck, 1979; Beck & Clement, 1991) produced improvements in elementary students’ standard achievement test scores of between 20 and 40 percentile points over a 3-year period. The intervention was the addition of only 30 min per day of timed practice and charting to an otherwise typical elementary school curriculum. Binder and Bloom (1989) described fluency-based corporate training programs that produced new sales trainees considered by their management to be more knowledgeable than senior sales representatives with up to 6 years of experience. Johnson and Layng (1992) reported results of a fluency-based adult literacy training program that were greater in magnitude than those produced by any other program funded by the Job Training Partnership Act. In the same publication they cited comparably superior results with children at the Morning-side Academy in Seattle and with precollege students at Malcolm X College in Chicago. The size of these effects suggests that fluency-based instruction may offer a cost-effective weapon against the increasingly acknowledged failure of the American education system. If confirmed by further systematic research, these results may lead to a fundamental shift in our understanding and design of optimally effective instructional programming—taking fluency into account.

The work on fluency has combined formal research with extensive field in-
vestigation and development conducted in demonstration programs, plus application in hundreds of precision teaching classrooms since the mid-1960s. Most of this work has not been documented in the scientific literature, but many of the empirical generalizations derived by fluency researchers and practitioners over the last 30 years suggest opportunities for important systematic research.

This article is intended to fill important gaps in the conceptual and historical record so that future researchers and practitioners can work from a full appreciation of what has come before, and make contact with current and past contributors. It brings together an extensive list of references on the topic, and provides context and background commentary to support further investigation and discussion among interested readers.

DEFINITIONS OF FLUENCY

An advantage of the term fluency is that many people already understand it intuitively or metaphorically. This familiarity may arise from common use of the term with reference to language (as in “he speaks French fluently”). I have often begun corporate seminars, graduate classes, and teacher workshops by asking the audience “What is behavioral fluency?” prior to any explanation of the concept. Responses from participants virtually always reflect prior understanding of the term and its implications. For example, when asked to list associations with the phrase behavioral fluency, one group produced responses that included easy to do, mastery, really knows it, flexible, smooth, remembered, can apply, no mistakes, quick, without thinking, automatic, can use it, not tiring, expert, not just accurate, and confident. Each of these reflects one or more attributes of what we mean when we use this term to describe the goal of instructional programming.

As currently defined, fluency is the fluid combination of accuracy plus speed that characterizes competent performance (Binder, 1988b, 1990a). Fluency has also been described as a combination of quality plus pace (Haughton, 1980). Other terms equated with fluency are automatic (Haughton, 1972a) and second nature performance (Binder, 1990a). A plain-English description of fluency is that it is doing the right thing without hesitation (Binder, 1988b).

The features ascribed to fluent performance closely resemble those traditionally associated with mastery. In defining the desired outcome of instruction, Barrett (1977a) explained that “Stability or predictability of performance is, then, vital in defining skill mastery” (p. 183). Gagne’s descriptions of mastery as “immediately accessible” and “performed with perfect confidence” (Gagne, 1970, 1974; Gagne & Briggs, 1974) have had significant influence on fluency researchers since the 1970s. In the final analysis, the term fluency is a metaphor reflecting all of these qualities, referring to a collection of observations about relations between response frequency and critical learning outcomes.

The empirical definition of fluency is related to its measured effects. When learners achieve certain frequencies of accurate performance they seem to retain and maintain what they have learned (Berquam, 1981; Kelly, 1995; Orgel, 1984); remain on task or endure for sufficient periods of time to meet

---

1 The term retention refers to the relation between behavior frequencies at two points in time, between which the individual has had no opportunity to emit the behavior. Maintenance, on the other hand, refers to the relation between a behavior’s frequency at two points in time, between which the individual has an opportunity to emit the behavior to produce reinforcement in the natural environment. It is an empirical question as to whether the frequency required to make a behavior “useful”—capable of being emitted, reinforced, and thereby maintained in its natural environment—is the same as the frequency that will ensure retention of the behavior after a period of time in which it has not occurred.
real-world requirements, even in the face of distraction (Binder, 1984; Binder, Haughton, & Van Eyk, 1990; Cohen, Gentry, Hulten, & Martin, 1972); and apply, adapt, or combine what they learned in new situations, in some cases without explicit instruction (Binder, 1976, 1979d, 1993a; Binder & Bloom, 1989; Haughton, 1972a; Johnson & Layng, 1992, 1994). When a combination of accuracy plus speed of performance optimizes these outcomes with respect to a specific behavior class, that is the level of performance that has been defined as “true mastery” of the behavior (Binder, 1987). Haughton (1980) captured this definition in an acronym by specifying what he called retention-endurance-application performance standards, or REAPS.

A NEW PARADIGM?

I have previously suggested that fluency represents a new paradigm in the analysis of complex behavior and the design of instruction (Binder, 1993a; Pennypacker & Binder, 1992). Although the term may be overused, it seems appropriate in this case. In his historic work, The Structure of Scientific Revolutions, Kuhn (1970, pp. 10–11) used the term paradigm to refer to developments in scientific method and practice that “attract an enduring group of adherents” and that are “sufficiently open-ended to leave all sorts of problems for the redefined group of practitioners to resolve.” Because developments associated with fluency have produced discontinuous changes in practice among a community of researchers and practitioners with respect to the definition of instructional outcomes and the measurement of instructional effectiveness, in the design and implementation of instruction, and in efforts to account for and reverse educational failure, they arguably represent a ground-shifting development worthy of this term. Despite the fact that the measures and methods of fluency initially evolved from past work in operant conditioning, their implications have subsequently led in directions that are truly revolutionary and unlike what preceded them. The remainder of this article is devoted to description of related historical developments and explication of their practical and scientific ramifications.

EARLY HISTORICAL DEVELOPMENTS

Origins in Free-Operant Conditioning

The work in behavioral fluency tracing its origins to free-operant conditioning insofar as fluency researchers and practitioners have explicitly studied and tried to produce streams of continuous responding rather than paced or controlled opportunities to respond (Barrett, 1977b; Binder, 1978b, 1993a; Lindsley, 1964, 1972, 1996a).

Skinner’s (1938) continuous measurement of behavior frequency in operant conditioning experiments revolutionized the study of behavior (Bjork, 1993, p. 93ff.). He observed later in his career that response frequency measures and the cumulative response recorder may have been his most important contributions (Skinner, 1976). Indeed, virtually all of the basic discoveries made in the research laboratories of Skinner, his students, and colleagues involved single-subject designs with continuous recording of free-operant response frequencies on cumulative recorders. In contrast to traditional estimates of response probability based on percentage correct calculations, Skinner (1938) pursued a program of research in which “rate of responding is the principal measurement of the strength of an operant” and where “probability of action has been attacked experimentally by studying the repeated appearance of an act during an appreciable period of time” (Skinner, 1953, p. 70). The glossary in Schedules of Reinforcement (Ferster & Skinner, 1957) defines probability of response as “the probability that a response will be emitted within a specified interval, inferred from its observed frequency under comparable condi-
tions” (p. 731) and strength of response as “sometimes used to designate probability or rate of responding” (p. 733).

Despite these seemingly fundamental views concerning the importance of behavior frequency, when Skinner and his colleagues began research in programmed instruction, an effort to extend basic laboratory discoveries into education and training, they generally dropped response frequency measures in favor of more conventional percentage correct or accuracy-only assessments (Skinner, 1954, 1968). In retrospect, this may be why fluency is only now emerging as a key element in the design of behavioral instruction: Most behavioral educators abandoned the frequency measure, except occasionally when monitoring problem behavior, more than 30 years ago.

**Precision Teaching and the Standard Celeration Chart**

Ogden Lindsley took exception to the trend away from frequency measures in educational applications. During the 1950s and early 1960s, Lindsley worked with Skinner directing the first operant conditioning laboratory for humans in which he confirmed and extended principles and procedures, originally developed in the animal laboratory, to human behavior, and coined the term behavior therapy as a way of distinguishing applied operant conditioning from psychotherapy (Lindsley & Skinner, 1954; Skinner, Solomon, & Lindsley, 1954). As in the animal laboratory, Lindsley relied on cumulative response records of behavior frequencies as the basic measurement and analysis technology—often simultaneously monitoring multiple operators with separate cumulative recorders.

During the early 1960s, Lindsley and his associates (prominently B. H. Barrett) applied functional behavior analysis in the laboratory to the diagnosis and remediation of retarded behavior (Barrett, 1965, 1969, 1971; Barrett & Lindsley, 1962; Lindsley, 1964). This work led to the development of precision teaching (Binder, 1988b; Binder & Watkins, 1990; Kunzelmann, Cohen, Hulten, Martin, & Mingo, 1970; Lindsley, 1972, 1990; White & Haring, 1976), in which teachers and their students used behavior frequency measures and the standard behavior chart (Pennypacker, Koenig, & Lindsley, 1972) to monitor individual classroom programs and make educational decisions.

Vargas, participating in both the broader tradition of behavioral education and in the subcommunity of precision teachers, wrote that

> Teaching is not only producing new behavior, it is also changing the likelihood that a student will respond in a certain way. Since we cannot see a likelihood, we look instead at how frequently a student does something. We see how fast he can add. The student who does problems correctly at a higher rate is said to know addition facts better than one who does them at a lower rate. (1977, p. 62)

This statement, rare among mainstream behavioral educators, eloquently repositions behavior frequency at the heart of behavioral instruction.

The standard behavior chart (more recently known as the standard celeration chart; see Figure 1) provided a measurement advance comparable to the cumulative recorder. Initially, Lindsley created the standard chart so that teachers sharing graphs of behavior frequencies would be able to share data more efficiently, based on a standard “graphic language.” By allowing students, teachers, and researchers to monitor behavior frequencies in a standardized graphic format, this tool reduced the time required to share data sets in a group from 20 to 30 min to about 2 to 3 min per chart (Lindsley, 1971).

An important feature of the standard chart is its combination of a linear abscissa for calendar time with a logarithmic ordinate for behavior frequency. The log scale was originally used to compress an entire range of human frequencies (from one per minute to one per day) onto a single graph. Lind-
sley and his associates soon discovered, however, that the semilogarithmic graphic space transforms learning curves into projectible straight-line trends (Koenig, 1972; Tukey, 1977) and allows calculations and projections of celeration, the first easy-to-quantify and visualize measure of learning rate in the literature. Celeration (either acceleration or deceleration) is the trend in a time series of frequencies expressed as a multiplication or division in frequency per week of calendar time. Celeration quantifies rate of change in frequency. For example, a trend that doubles a behavior frequency in a week (and, it so happens, is parallel to a line going corner-to-corner on the standard chart) is called ×2.0 celeration per week, and one that divides average frequency by 3.0 in a week is called a ÷3.0 celeration. ×1.0 is a flat line, with no trend (Johnston & Pennypacker, 1980; Pennypacker et al., 1972). On a semilogarithmic chart, the visual angle of a given celeration is the same, independent of the frequency at which it begins. For example, a celeration doubling (×2.0) from one per minute to two per minute in a week forms the same angle with the horizontal as a celeration doubling from 60 per minute to 120 per minute or from 150 per minute to 300 per minute in a week. Decelerating from 100 per minute to 25 per minute (÷4.0) is the same as from four per minute to one per minute, and so on. By representing both frequency and celeration in standard graphic and quantitative units, the standard chart clearly differentiates between changes in performance levels (frequencies) and changes in learning rates (celerations) (Lindsay, 1996b).

Precision teachers learned to use projected celerations (later called cel-
eration aims) to set minimum acceptable learning rates (Koenig, 1972; White & Haring, 1976) for daily or weekly instructional decision making. As long as the actual data did not fall below the projected celeration line for more than 2 days in a row, the program continued. Data failing to accelerate or decelerate as rapidly as the celeration aim for several days in a row prompted a change in the program. Analogous in use to the within-session cumulative response record in the laboratory, the standard chart became an ongoing decision-making tool for practitioners and behavior scientists studying changes in frequencies across sessions. It allowed easy inspection, quantification, and decision making based on the next derivative of behavior frequency, change in daily frequency per week (Kazdin, 1976).

Lindsley's goal (1972) was to put scientific methods in the hands of teachers and students—to transform classrooms into places for data-based discovery, fully integrated with educational practice. Adapting the laboratory model of direct continuous recording, Lindsley and his associates timed and counted various types of classroom behavior for extended periods of time during the early years of their work in education. They began precision teaching by transferring laboratory strategies and tactics into the classroom, using the standard behavior chart to monitor and analyze performance and learning. In fact, early students of Lindsley studied many of the response classes and phenomena addressed by other applied behavior analysts.

For example, Kunzelmann (1965) completed a master's thesis with Lindsley by designing a transducer for monitoring frequencies of out-of-seat behavior in the classroom. Haughton's (1967) doctoral dissertation likewise dealt with the relatively traditional behavioral topic of reinforcer sampling, presenting data on a precursor of the standard behavior chart.

Initially, precision teachers measured how much time students required to complete practice sheets and calculated count per minute with a fixed numerator and variable denominator (Lindsley, personal communication, 1995). After a while, the practice of collecting brief (e.g., 1-min) fixed samples of behavior frequencies emerged as a critical component of precision teaching (Haughton, 1972a; Kunzelmann et al., 1970; Starlin, 1972), in part for calculation convenience. Although Lindsley (personal communication, 1995) at first resisted short measurement intervals, preferring to record behavior over extended periods of time as in the operant conditioning laboratory, proponents of brief timings persevered. They quickly recognized the sensitivity of brief timings to differences in skilled performance, and began to use brief timings as a rapid and inexpensive method for gathering descriptive information about various types of human behavior. This methodological shift toward using brief fixed timings to calculate behavior frequencies led to initial discoveries about fluency among precision teachers (Haughton, 1972b; Kunzelmann et al., 1970).

Professional Communication Based on Charts Rather Than Publications

Those involved in precision teaching did not seek to publish in the way that is generally maintained by academic contingencies of reinforcement. There seem to have been three primary reasons for this turn of events. First, most were practitioners who did not pursue publication for career advancement. Second, discoveries in precision teaching were progressing more rapidly than journal or book publication cycles could match, and this discouraged even the academics among precision teachers from formally reporting findings or practices that would be obsolete by the time of publication. Third, from his own extensive history of publications in human operant conditioning, Lindsley (personal communication, 1974) concluded that publications did not
change professional behavior sufficiently to justify the effort required for publishing in academic journals. He consequently discouraged early precision teachers from devoting time to traditional publications for professional communication. Therefore, the discoveries of precision teaching remain comparatively undocumented in the academic literature (Lindsley, 1990).

A few years after the inception of precision teaching, Lindsley and his associates started the Behavior Bank (Koenig, 1971; Lindsley, Koenig, Nichol, Kanter, & Young, 1971), a computerized database into which practitioners deposited intervention data summarized from standard behavior charts as frequencies, calendar durations, and celerations. Originators of the Behavior Bank planned that precision teachers would accumulate inductive research data and would maintain their scientific communication via access to this common database and by sharing standard behavior charts (as was the practice with cumulative records during the early days of operant conditioning). The Behavior Bank was a technology before its time, prior to the advent of personal computers and dial-in networks, and died within a few years, although Lindsley (personal communication, 1995) still maintains data from thousands of chart projects stored on magnetic tapes.

During the 1970s a few precision teaching textbooks appeared (Kunzelmann et al., 1970; White & Haring, 1976). In conjunction with open monthly chart-sharing sessions held at Barrett’s Behavior Prosthesis Laboratory, Binder published the Data-Sharing Newsletter from 1977 to 1983 (to be republished by PT/MS, Inc., P.O. Box 95009, Nonantum, MA 02195), which informally reported data sets and discoveries, large and small, among several hundred practitioners and researchers. McGreevy began the Journal of Precision Teaching in 1980 (now edited by McDade at Jacksonville State University).

Like resistance to publication of cumulative records by nonbehavioral journals before inception of the Journal of the Experimental Analysis of Behavior, mainstream behavioral journals refused for many years to publish data displayed on standard behavior charts. Thus, precision teaching and its discoveries have remained more an oral than a written tradition in the field of behavior analysis, based on the personal exchange of charted data from many thousands of single-subject classroom interventions and on charts presented at professional conferences. This article, and other recently published papers (Binder, 1988b, 1992a; Binder & Watkins, 1990; Eshleman, 1990; Lindsley, 1990, 1991, 1992, 1994, in press; Potts, Eshleman, & Cooper, 1993) seek to reverse that trend, and to encourage formal research and publication of results.

The volume of data accumulated by precision teachers, although not shared widely, is nonetheless remarkable. For those who suggest that precision teaching data do not comprise a scientifically valid body of findings or are merely correlational in nature, it is worth recalling the early history of operant conditioning. For over 25 years, without a journal of their own, operant conditioners shared sets of single-subject replications via collections of cumulative records. In precision teaching, early reports of findings reflect a similar strategy of accumulated multiple baseline replications across subjects and response classes. For example, Starlin’s (1971) earliest published analyses of reading proficiency and of the component behavior frequencies required to achieve reading competence were based on several hundred individual replications across students. Although many of the reported discoveries of precision teaching certainly should be subjected to controlled studies of a more traditional nature, the number of replications upon which these claims are based far surpasses the quantities of data involved in most contemporary dissertations or published behavioral studies. I hope that the tradition that has evolved from this informal com-
munication network will help to guide more formal research in the future among those for whom such research is reinforced.

**KEY DEVELOPMENT: FREQUENCY AIMS**

*Accuracy is Not a Sufficient Criterion for Mastery*

Eric Haughton was one of Lindsley’s first precision teaching doctoral students. During the late 1960s, Haughton (1972a) and his associates observed that the mere presence or accuracy of a response class in the repertoire of a learner is not sufficient to ensure progress through a curriculum sequence that depends on that response class as a prerequisite or component. They found, for example, that if students were not able to write digits or read random digits at around 100 per minute, they would not be able to progress smoothly through acquisition and mastery of computational arithmetic (Haughton, 1972a; Starlin, 1972). Yet with daily practice on these elementary skills (originally called *tool skills*), students were able to achieve higher performance frequencies that, in turn, enabled them to acquire and develop useful frequencies of computation (50 to 60 per minute) and to progress successfully through the math curriculum. They extended this discovery to writing, reading, and spelling curricula as well (Haughton, 1972a; Starlin, 1971; Starlin & Starlin, 1973a, 1973b, 1973c, 1973d).

Haughton (1972a) wrote that with respect to academic tool skills such as writing digits 0 to 9, reading random digits, or saying the sounds for letters, “aims between 100 and 200 movements per minute indicate proficient performance, whatever the curriculum area” (p. 32). At the same time, he and his associates found that although errors may be difficult to correct when overall response frequencies are low (e.g., reading below 50 words per minute), errors became easier to decelerate when overall performance was at higher frequencies (e.g., above 50 or 60 words per minute) (Haughton, 1972a). This finding foreshadowed Haughton’s (1980) later guideline that only when students can perform at approximately half the proficiency level for a given skill are they most likely to engage in and profit from independent practice.

Confirmed in many ways since (Binder & Bloom, 1989; Evans & Evans, 1985; Johnson & Layng, 1994; Lindsley, 1992), this principle of *minimum component behavior frequencies* became an underpinning of fluency-based instruction and set the stage for significant improvements in the efficiency of instructional programming (Beck & Clement, 1991; Binder & Watkins, 1990; Johnson & Layng, 1992). What many educators assumed to be “learning disabilities” or “learning problems” seemed to wane when students were allowed and encouraged to practice key components of complex behavior to the point at which they could perform each component at relatively high frequencies (Beck, 1979; Binder, 1991b; Haughton, 1972a; Johnson & Layng, 1992). These observations began to make clear that achieving a high performance frequency increases the range of a student’s potential performance capacity, enabling that individual to meet any performance requirements at or below the attained level (Elizabeth Haughton, personal communication, 1995). This was a radically new idea for precision teachers in the late 1960s.

**Constraint on Reinforcement Effects**

These observations revealed a constraint on the ability of reinforcement to increase the frequency of composite behavior. When Haughton (1972a) and his associates first began to recognize the importance of behavior frequencies as indicators of skill proficiency, they attempted to reinforce performance of basic academic skills. But low frequencies of tool skills (e.g., writing digits) imposed ceilings on the acceleration of composite behavior frequencies (e.g., writing answers to math problems),
and previously identified reinforcees alone proved incapable of increasing frequencies of the composites to the desired levels. Only prompting and reinforcing performance of components led to higher composite frequencies. Thus, new observations about response–response frequency relations revealed a previously unrecognized constraint on the potential of reinforcement procedures to increase frequencies of complex behavior. Even ordinarily strong reinforcement contingencies, identified separately with other response classes in the same individual, might prove to be ineffective if applied to composite behavior when component behavior frequencies are low. This finding also led to research designs in which experimenters must be certain beforehand that component behavior frequencies do not artifically constrain the growth of composite responses being subjected to experimental procedures designed to increase their frequencies (Binder, 1984).

Programming Based on Component–Composite Relations

Initial use of performance aims focused on tool skills related to reading, writing, and computational math. An understanding of the relations among tool skills and basic academic skills led Haughton to use a chemical analogy, referring to a general relation among response classes as elements and compounds (Haughton, 1981a). His analogy suggested that, like atoms requiring a certain valence or energy to combine, behavioral elements require a certain frequency to form compound response classes. Others (Barrett, 1977a; Binder, 1978a), borrowing from the literature of perceptual-motor learning (e.g., Gagne, Baker, & Foster, 1950), first used the terms component and composite to refer to this general part–whole relation as applied in precision teaching.

Curriculum analyses and designs during the 1970s and early 1980s focused on identifying relations between behavior components and composite repertoires. Haughton (1972a) studied correlations in log-log scatter plots between frequencies of components and composites in the repertoires of individuals and groups. Initial functional analyses studied component–composite relations by attempting to build frequencies in components and then observing the effects on composites (Haughton, 1972a). Van Houten (1980, pp. 24–25) described a procedure that used the frequency of writing answers to long multiplication and division problems (composite) as a dependent variable to assess the effects of increasing frequencies of writing answers to basic multiplication facts (components).

Extending the approach beyond academic behavior, Haughton and his associates worked with teachers of multiply disabled students who exhibited severe deficits in fine and gross motor control. Collaborating with Mary Kovacs, who was trained as a physical therapist and nurse (Haughton & Kovacs, 1977; Kovacs & Haughton, 1978), and with Anne Desjardins and Bev Palmer (Binder, 1979a), Haughton identified a set of fundamental component skills, originally called The Big 6 (reach, point, touch, grasp, place, release) and later enlarged to The Big 6 Plus 6 (including twist, pull, push, tap, squeeze, shake). They also developed a taxonomy of behavior components involving gross motor control of trunk, arms, legs, and head (Kovacs, 1978). Estimating competent performance ranges using brief timed samples of adult performance to establish aims and providing isolated practice with these fine and gross motor skill elements, Haughton and his associates enabled severely disabled people to achieve previously unattainable functional skills (Binder, 1991b). Binder and his associates extended this work to multidisciplinary programming with physical, occupational, and language therapists (Binder, 1981a, 1981b; Binder & Pollard, 1982; Burgoyne, 1982;
Imbriglio, 1992; Pollard & Binder, 1983).

Perhaps the most dramatic success story during these years was the case of Terry Harris, a boy born with severe cerebral palsy and diagnosed as likely to be institutionalized, nonverbal, and nonambulatory. Eric and Elizabeth Haughton worked with Terry and his parents from early childhood (Binder, 1991b). Today, in his 20s, Terry attends graduate school, drives, skis, and is a motivational speaker, despite the persistence of his neuromuscular handicap. His success was built on many thousands of hours of practice to achieve fluency on the most basic fine and gross motor elements and an entire repertoire constructed of those elements, using precision teaching methods in a progressive curriculum of component–composite relations. (Records of this case include a videotaped presentation from the 1990 International Precision Teaching Conference featuring Terry, his mother, and Elizabeth Haughton, his teacher, in addition to charted data.)

Much work at Barrett’s Behavior Prosthesis Laboratory and associated agencies (see below) during the late 1970s focused on application of these principles to a broad range of self-care and vocational skills among the severely disabled, especially development of materials and procedures for assessing and practicing components in isolation prior to combining them into chains (Barrett, 1977b, 1979; Binder, 1976; Bourie & Binder, 1980; Pollard, 1979; Solsten & McManus, 1979). These procedures provided alternatives to accuracy-based backward chaining methods that had proven to be unreliable in producing lasting, functional repertoires for many disabled learners (Barrett, 1977a).

FROM AIMS TO REAPS

Seeking Performance Standards

As Haughton and his associates worked to identify performance aims, they frequently found it necessary to raise what they had thought to be appropriate criteria to higher levels, because students were able to achieve them and because achieving more rapid performance of components usually led to easier learning and better performance of composites. For example, Haughton (1972a) reported that reading orally at 100 words per minute and writing answers to basic arithmetic problems at 40 to 50 problems per minute were sufficient to ensure subsequent progress through curriculum. By the end of the 1970s, commonly used aims for those skills were 250+ words per minute (Starlin, 1979) and 80 to 110 problems per minute (Haughton, 1980), respectively. Acknowledging this evolving development of fluency standards, every list of performance aims distributed by Haughton included a revision date set 1 year after the date of creation, indicating that the aims recommended in any given document should be reviewed at least once per year, to see if they reflect current evidence.

During that period, some precision teachers had begun to set aims with their students using levels of performance significantly below normal adult frequencies (Howell & Kaplan, 1979; White & Haring, 1976). In fact, some practitioners even suggested lowering aims to account for age and level of disability. An educational practice known as curriculum-based measurement (Binder, 1990b; Deno, 1985) was influenced by precision teaching work conducted in Minnesota by Clay and Ann Starlin (1973a). This approach reduced the notion of competency-based aims to norm-based criteria, however, using class averages as performance standards instead of criteria intended to reflect empirically determined competence levels and to ensure successful learning and application. The use of “handicapped” aims and of class averages to set aims contains an inherent flaw, if the objective is to produce competent performers. When applied in schools in which classroom medians fall far below levels shown to represent...
competence in the community (e.g., Wood, Burke, Kunzelmann, & Koenig, 1978), these approaches virtually institutionalize incompetence in the form of suboptimal performance criteria. The general practice of setting educational goals based on norms rather than on empirically validated measures of competence may be responsible for the increasing prevalence of illiteracy and other skill deficits within the schoolgraduate population. Haughton and his colleagues pushed in the opposite direction, establishing aims by collecting measures of competent adult performance, and encouraging students to achieve their "personal best" levels for every skill.

Setting Aims Using Frequency Sampling

Wood et al. (1978) collected brief frequency samples of math skills performed at peak levels by high-performing and low-performing eighth graders as well as by professionals who used arithmetic in their jobs. The data revealed that adult professionals were generally higher in performance frequency than eighth graders at the top of their classes, except in skills seldom used by adult professionals (e.g., fractions and decimal arithmetic). Barrett (1979) made similar comparisons of performance on 16 prevocational and preacademic skills among competent adults, normal children, and institutionalized disabled students in her laboratory classroom. Although all performed at 100% accuracy and were therefore indistinguishable from one another on an accuracy scale, the ranges of behavior frequencies for each population clearly separated competent adults from normal children and distinguished both groups from the disabled students.

The approach of sampling performance of various populations introduced an important element of naturalistic observation that would have been impossible with accuracy-only metrics. As a rule of thumb, on any well-prac-

ticed skill in a homogeneous adult population, the range of frequencies represented by as few as a half dozen individuals generally provides a reasonable estimate of performance levels in a larger population. (To convince yourself, ask a half dozen competent adults to write answers to simple addition problems on a sheet containing 120 or more such problems for 1 min as rapidly as possible. You will likely find that most of the individuals will write between 80 and 110 answers per minute.) Such an empirically determined range of behavior frequencies is quite different from an arbitrarily chosen percentage correct criterion. Unlike percentage correct, a dimensionless quantity (Johnston & Penypacker, 1980), behavior frequency is a standard unit of measurement and places frequency-based instructional design and assessment squarely in the domain of natural science (Barrett, 1977a, 1979; Binder, 1995). For well-practiced behavior in a normal adult repertoire, samples of competent adult performance generally provide a good first approximation for setting instructional aims. Prior to completion of controlled studies designed to identify optimal performance aims for specific skills, behavior frequency sampling methods (sometimes known as snapshots among precision teachers) provide important tools for instructional designers and practitioners.

REAPS: Aims Based on Critical Learning Outcomes

During the late 1970s, Haughton initiated use of the term REAPS (retention/application performance standards), suggesting that we set aims empirically by determining what levels of performance ensure retention and application of skills (Haughton, 1981b). Shortly thereafter, the term expanded to REAPS (retention-endurance-application performance standards), reflecting observations that achieving high performance frequencies seemed to increase the likelihood that students
would maintain attention to task over extended durations of performance and in the face of distraction—what he and others called the *endurance* of performance (Binder, 1984; Binder et al., 1990; Cohen et al., 1972; Haughton, 1980). Endurance became a new subject for instructional research. The REAPS acronym set a long-term research agenda aimed at determining, for every response class of interest, performance standards that ensure these critical learning outcomes.

**Evidence to support REAPS**

The determination of performance standards based on the criterion that they optimally support retention, endurance, and application suggests a virtually endless program of investigation that could keep researchers busy for decades. To meet the challenge posed by Haughton’s acronym, we would need to determine, for each behavior class, the frequency ranges required for optimally supporting each of these outcomes. Moreover, the frequencies are likely to vary for any given class of behavior. For example, an individual might permanently retain or remember basic math facts practiced to 60 or 70 per minute, with negligible improvements in retention beyond that range, yet continue to improve in the ability to apply the skill in mental math as it accelerates beyond 100 per minute. That is, the optimal frequency for retention might be different from that for endurance or application. Multiplied by the total number of response classes in a human repertoire, this challenge may be practically impossible to address for every important one. Nonetheless, practitioners and researchers will continue to investigate and experiment with levels of performance and their effects in several important domains, most notably the basic academic and intellectual skills.

Simply demonstrating in a systematic fashion that higher performance frequencies improve outcomes in one or more of the three categories for any behavior class is itself a notable accomplishment, one that can surely inspire many theses and dissertations in the future. What follows is a brief summary of some key findings related to each of these outcomes, most of which beg for replication and systematic experimental analysis.

**Retention.** A variety of classroom instructional design projects have demonstrated effects of frequency building on retention. Disabled students who had previously failed to acquire or maintain behavior chains (e.g., assembly or dressing skills) with standard accuracy-based backward chaining procedures were able to combine and apply behavior components in chains after repeated daily practice of each component in isolation had increased performance frequencies (e.g., Pollard, 1979; Solsten & McManus, 1979). Although these projects were clinical in nature and did not involve formal control conditions, they were essentially multiple baseline replications across individuals. They are generally referred to as support for the application aspect of REAPS. However, many teachers of the disabled have worked with students who do not retain components of even the simplest chains for more than a few hours or days after accuracy-only chaining procedures. The results of these programs suggested that increasing behavior frequencies improves retention, in the sense that retention of components is a minimal prerequisite for subsequently integrating them into chains.

In addition, college students who practiced calculus formulas and rules using timed flash cards to achieve aims of saying 50+ facts per minute were able to perform nearly twice as accurately on tests 6 weeks later as those who did not achieve high frequencies (Orgel, 1984). Berquam (1981) demonstrated similar relations between retention and performance frequency. Kelly (1995) used a within-subject yoked design to separate the effect of mere repetition from that of achieving more rapid responding, and supported
the conclusion that achieving more rapid performance yields greater retention.

Endurance. Binder (1982, 1984; Binder et al., 1990) has reported research on the ability of students to perform for extended periods of time as a function of initial performance frequency. Early observations with disabled students demonstrated that for those with low levels of performance, practice durations as short as 3 to 5 min were too long to sustain steady performance, even with added reinforcement procedures. Students slowed their performance within the first minute or two, and often exhibited off-task or disruptive responses. When required performance durations were shortened to 1 min or less, performance frequency jumped or turned up and exhibited less variability, and students stopped emitting off-task behavior. Changing performance durations affected frequencies of correct and error performance as well as celerations. Working for shorter intervals often enabled students to achieve high levels of performance faster.

These effects are easy to observe in any population in which individuals have not yet achieved competent levels of performance. Application of these findings to instructional programming involves working with very short intervals (e.g., 10 s) called sprints (Haughton, 1980) until students are able to achieve aims, then gradually lengthening practice intervals to build endurance (Bourie, 1980; Desjardins, 1981). Haughton, Maloney, and Desjardins (1980) adapted the count per minute standard celeration chart for such procedures by changing the daylines into successive minute-lines for charting repeated sprints. Johnson and Layng (1994) have reported using a version of this methodology in the Morningside model.

Johnson (personal communication, 1996) reports a cautionary note that students who achieve high frequencies for brief durations within sessions, without continuing on successive days to practice until they achieve aims for longer durations, may not exhibit the same degree of retention or application during later sessions as if they had been required to achieve aims for longer durations on successive days. This finding emphasizes the importance of distributing practice over multiple sessions, and of checking performance frequencies on more than one day to be certain they are retained.

Two unpublished sets of pilot data obtained by the author provide templates for future endurance research. In the first (Binder, 1984), teachers collected samples from 75 students repeatedly writing digits 0 to 9 for varying durations, once per day, in ascending sequence: 15 s, 30 s, 1 min, 2 min, 4 min, 8 min, and 16 min. The distribution of performances across the population for the 15-s interval ranged from less than 20 per minute to over 150 per minute. Each subject's median count per minute across all durations placed him or her in a frequency bin, each bin spanning a range of 20 per minute. Figure 2 summarizes the results, each data point representing a median frequency at a given duration for the individuals in a given frequency bin. These data show greater performance decrements at the long intervals for subjects with lower performance frequencies. Around 70 per minute appears to be a cut-off point beyond which higher initial frequencies do not predict greater ability to sustain prolonged performance. Using this approach (being sure to study at least an order of magnitude range in both behavior frequencies and performance durations), future investigators may be able to identify such cut-off points for other types of behavior.

The second pilot design (Binder, 1979c) is a free-operant analogue of automaticity experiments conducted by cognitive psychologists (LaBerge & Samuels, 1994) who used latency measures in trials procedures. Two adult subjects performed five different tasks in successive 3-min intervals: reading numbers, saying answers to simple ad-
Figure 2. Points represent group median count per minute at each performance duration for each of eight groups of subjects. Each group contained subjects whose median performances across all durations were within the indicated frequency range. \( N = 75 \).

dition problems (sums to 18), reading printed anglicized names of Hebrew characters, saying numbers in response to the names of Hebrew characters (previously learned in a paired associate procedure), and adding Hebrew characters by using the previously learned paired associate to assign numbers to the characters (an example of stimulus equivalence). Subjects performed all tasks by reading aloud from practice sheets into a microphone attached to a voice-operated relay with electromechanical equipment for counting and recording responses on a cumulative recorder. Figure 3 shows pairs of cumulative records for each task, each pair representing the performance of the 2 subjects during a single session. Note that the 2 subjects perform the first three tasks at about the same frequencies, as would be expected because these three are well-established arithmetic and reading skills found in competent adults. On the fourth task, a newly learned paired associate, the 1st subject, who had completed more practice sessions, performed at a higher frequen-
cy than the 2nd subject. And on the fifth task, which required the newly learned paired associate as a component, the 1st subject performed considerably more rapidly, as would be expected. After a brief rest period, both subjects repeated the same tasks, this time wearing headphones through which they heard random numbers (a distracting stimulus) for 30-s periods halfway through each session. Figure 4 shows cumulative records of these performances, with sup-

Figure 3. Each pair of cumulative records represents the pair of subjects performing the listed behaviors, recorded by means of a voice-operated relay.

Figure 4. Each pair of cumulative records represents the same pair of subjects as in Figure 3 performing the listed behaviors, recorded by means of a voice-operated relay. Tick marks indicate onset and termination of a distracting auditory stimulus associated with the suppression ratios.
pression ratios calculated as frequencies between the tick marks divided by frequencies averaged for the periods before and after the marks (Estes & Skinner, 1941). These suppression ratios and corresponding visual dips in the records between tick marks reflect proportional decrements in responding associated with the distracting stimulus.

Although others will surely need to replicate this experiment to test the findings further and apply the design with other response classes, these pilot data indicate that lower performance frequencies may be associated with greater distractibility, measured as relative suppression of responding during presentation of an external stimulus. This model applies free-operant laboratory methods to measure distractibility as an alternative to the more cumbersome and less sensitive latency-based trials procedures generally used by cognitive researchers and by some behavior analysts.

In combination, these two pilot studies reflect expected characteristics connoted by the term endurance: the ability to continue performing over increasing durations and in the face of environmental distraction (much like the strong long-distance runner who can persist without stumbling, even when encountering obstacles in the path). Most important, they offer designs for further analysis.

Application. By far the greatest amount of evidence exists to support the conclusion that increased performance frequencies improve application. By application we mean integration of component response classes into composite response classes. Haughton’s (1972a) original report indicated that increasing the frequencies of component skills supports more rapid learning and performance of composites. This basic finding has been replicated countless times in precision teaching classrooms for regular or mildly disabled students (Beck, 1979; Evans & Evans, 1985; Johnson & Layng, 1992; Lindsley, 1992; Maloney, Desjardins, & Broad, 1990; Mercer, Mercer, & Evans, 1982; Starlin, 1972; Van Houten, 1980). Classroom and vocational projects with severely disabled learners (Binder, 1976, 1979d; Pollard, 1979; Solsten & McManus, 1979) demonstrated that frequency building of components not only allows more rapid acquisition of composites but sometimes seems to produce composites with virtually no formal instruction—an effect that Johnson and Layng (1992) have called response aduction (Andronis, 1983; Epstein, 1985).

An earlier study by Binder (1979d) demonstrated the effects of building component frequencies in a free-operant analogue of Sidman’s (1971) mediated transfer procedure with 4 institutionalized disabled students. Subjects learned to read sight words corresponding to words in their existing speaking vocabularies (i.e., they could vocally name the actions and objects corresponding to the words and follow spoken directions using the words), achieving a 100% correct criterion at no higher than 12 words per minute. All subjects were then able to match printed words to objects and actions, and 2 were able to follow four-word written instructions (e.g., put ball in cup). After daily frequency-building procedures produced increases in oral reading performance frequencies, retesting revealed that all students increased frequencies of matching words to objects, and that all subjects could now follow written directions using the words, including the subjects who had previously been unable to do so after accuracy-only training on the sight words. Applying frequency-building methods to components in stimulus equivalence experiments provides a means for investigating the temporal dimension of intellectual skill.

Taken together, these observations point to a relation between performance frequency and application. Nonetheless, we need further controlled free-operant research to sort out the variables and to further define behavior frequencies associated with application in different skill areas.
Aims as Ranges of Performance

One final aspect of REAPS or aims is that they are best expressed as ranges rather than single frequencies (Haughton, 1980). This practice accounts for variation among individuals, analogous to normal ranges in medicine, as well as individual preferences. When given a choice, some students will practice until they can reach the highest possible level of performance, whereas others will settle for lower levels. A range allows for such individual variation, but the minimum criterion should nonetheless represent performance that will be retained, will endure, and can easily be applied.

TECHNICAL EVOLUTION

This section summarizes developments that have led to technical evolution in fluency research methods and educational practice. It focuses more on methods and induced principles than on the results themselves. In addition, by introducing key terms used by fluency researchers and practitioners, it prepares the reader for further investigation of that literature.

Stages of Learning

During the 1970s, precision teachers began to differentiate between stages of the learning process, separating the process for achieving accuracy from that for attaining fluency. For example, Binder (1976) referred to rate building as a stage beyond acquisition of accurate performance. White and Haring (1976), Haring (1977), and Haring and Liberty (1978) described stages of learning that include acquisition, fluency building, maintenance, application, and adaptation. Each stage involved different types of procedures and different criteria. Johnson and Layng (1992, 1994) have incorporated the implications of research on endurance (Binder, 1982; Binder et al., 1990) into the Morningside model, distinguishing among accuracy training, fluency building, endurance building, and applying.

An Emphasis on Practice

Practice is the repetition of a given response class after it has been accurately established in a repertoire. Based on the understanding that fluency is achieved with practice, Haughton (1980) recommended that at least half the time spent on education should be practice, with a complementary reduction in time spent on acquisition of new behavior. The rationale, like that suggested by Gilbert (1978), is that educational programs will be more effective in the long run if they produce a more focused, but truly mastered, repertoire rather than a broad but fragile repertoire. The latter might be said to characterize the usual educational approach in America, which introduces but never ensures mastery of a broad range of skills and knowledge.

Johnson and Layng (1992, 1994) have reported that when the basics are fluent, later learning becomes easier rather than more difficult (a topic rich with opportunities for controlled research). Thus, greater focused practice to achieve fluency in a foundation repertoire is more likely to be cost effective and time efficient than a broad accuracy-based approach to curriculum. This is in contrast to the typical instructional program in which most time is spent on acquisition or "establishing" the skills, with insufficient practice to ensure fluency.

Mainstream educators often characterize traditional drill and practice as outmoded, boring, and ineffective because it does not support the higher order problem-solving repertoires needed in today's world. Johnson and Layng's (1992) report that fluent prerequisites support easy acquisition of problem-solving repertoires contradicts this philosophical assertion.

Fluency-based instructional methods alter important features of traditional practice methodologies that make them ineffective and unpleasant (Binder,
First, much of traditional academic practice has no clearly defined objective other than to "get better." Whereas practice in such skills as martial arts or musical performance sets implicit time-based fluency aims as practice goals (smoothness, quickness), traditional accuracy-based educational assessment cannot measure improvement beyond 100% correct. In the absence of a fluency goal and feedback against that goal, practice receives little or no reinforcing consequences. With the addition of fluency aims and daily measurement, practice has a clearly defined goal, and advancement toward that goal might be reinforcing. Second, many practice procedures strain the endurance of individuals who have not yet achieved sufficiently high performance frequencies (Binder et al., 1990). Prolonged practice when behavior frequencies are low may be subjectively unpleasant and may occasion undesirable off-task behavior. Learners may not be able to maintain a given level of performance for more than a very brief interval (e.g., 15 or 30 s) when the performance level is far below its fluency aim. Brief practice intervals can provide a cost-effective antidote for these undesirable effects, often accelerating correct responding while reducing variability and problem behavior. Third, much traditional practice occurs under aversive control. Effective precision teaching methods positively reinforce improvement with feedback from charted data and encouragement from teachers or practice coaches.

Development in Instructional Design

Steps and slices. Precision teachers use the language of steps and slices to describe curriculum sequences (Starlin, 1972; White & Haring, 1976). A step is a phase change to a new class of behavior or a new subobjective (e.g., from writing digits 0 to 9 over and over again to writing digits as answers to math problems). To step back is to practice a prerequisite class of behavior as a form of remediation in a curriculum sequence. A slice is a subset of all possible instances of a particular behavior (e.g., writing answers to addition problems with sums to 10 as a subset of all possible addition problems with sums to 18). To slice back is to select a smaller set of behavior instances for practice to accelerate learning. The term curriculum weight (Haughton, 1979, personal communication) reflects an understanding that if learners are asked to practice too large a slice of behavior, the acceleration of that behavior may be "weighted down" by too heavy a burden—too many instances at once. These concepts, plus a variety of logical component—composite chunking and sequencing approaches to curriculum design, were the focus of precision teaching for many years (e.g., Howell & Kaplan, 1979; Starlin & Starlin, 1973b, 1973c, 1973d).

Learning channels. The concept and classification of learning channels represented an important advance in fluency-based curriculum analysis and design. Early precision teachers used verb channels (Kunzelmann et al., 1970) to indicate the type of movement represented by a given response class or pinpoint (Lindsay, 1972). The nomenclature of verb channels included such active verbs as say, write, touch, and mark. They provided unambiguous language for categorizing the form of behavior. Later, precision teachers combined inputs (antecedent stimuli) with verb channel outputs to describe behavior (e.g., oral reading is see/say words). Finally, Haughton (1980) introduced the learning channel matrix (Figure 5), a grid listing possible inputs down the left (see, hear, sniff, taste, touch, and free) with verbs across the bottom indicating actions (e.g., say, mark, type, write, tap, do, etc.). Many types of skills might fall into a given input/output combination or channel (e.g., see/say numbers, words, colors, and pictures of famous people).

Haughton (1980) used the term channel set to refer to the collection of
all skills in a given channel (corresponding to a cell in the learning channel matrix) and the performance standards associated with each. The channel matrix enabled curriculum designers to plan for a variety of different skills in the same curriculum area by viewing all possible input/output combinations on a single summary form. For example, see/say answers to math facts, hear/write answers, and hear/say answers might all be parts of a math curriculum. Each form of behavior could be assessed and practiced on its own or in combination with others. Lindsley (1992) cites evidence indicating that learning and performance in one channel are generally independent of (or cannot be predicted from) others, recommending explicit assessment and instruction in every channel of interest for a given curriculum area. Haughton (1977; Hastings County Board of Education, 1977a, 1977b) and his associates used the channel matrix to map out broad sets of curriculum in a range of different domains. Binder (1989) used channel language as a core component of the FluencyBuilding™ approach to instructional design, and Elizabeth Haughton (1993a, 1993b, 1994) has continued to apply this analytical framework in designing instructional activities and materials.

One implication of the learning channel matrix for setting aims is that the aim for a given form of behavior in one learning channel may be predictable from others in that same channel. For example, reading words at 150 to 250 per minute in the see/say channel helps to predict the pace at which an individual might be able to silently read the fronts of flashcards (McDade & Olander, 1990) and say the words written on the backs. Likewise, the frequencies of all see/write skills have predictable quantitative relations that
can be estimated with frequency sampling procedures in a given population. Categorizing types of behavior in this way may help to estimate appropriate performance aims for one behavior based on what is known of another in the same channel, without having to empirically establish REAPS for an infinite number of different specific types of behavior. For example, count per minute aims for a set of see/say biology flash cards will be approximately the same as for a set of history flash cards. On the other hand, some operators that one might naively believe would occur at the same frequencies do not (e.g., see/say numbers, see/say the colors of dots). So it's always best to confirm frequency ranges by sampling the performance of competent adults or experts.

Combining precision teaching with direct instruction. Maloney and Humphrey (1982) and Maloney et al. (1990) first combined the methods of precision teaching with Engelmann's direct instruction approach (Binder & Watkins, 1990; Engelmann & Carnine, 1982; Watkins, 1988). The intention was to optimize direct instruction's small-group teaching methods and empirically validated strategy-based instructional designs with the addition of the assessment and frequency-building methods of precision teaching. Johnson and Layng (1992, 1994) continued in this direction by adding significant refinements in instructional design and delivery methodology, influenced by the work of Tiemann and Markle (1990). They analyzed and sequenced curriculum to encourage generativity, the emergence of new behavior based on the principle of contingency addition (Andronis, 1983). Binder has applied similar analyses and design principles to develop curriculum and learning procedures for corporate sales professionals (Binder & Bloom, 1989).

Assessment Methodology

Placement in a curriculum sequence. Starlin and Starlin (1973a, 1973b, 1973c, 1973d) developed a precursor to what Johnson and Layng (1994) have recently called precision placement. By breaking curriculum sequences into fine steps and slices, and by sampling performance at points up and down those sequences, precision teachers place students in a curriculum based on performance frequencies.

Using ratios to predict performance. Precision teachers began to use ratios of component to composite behavior frequencies sampled in competent performers (e.g., writing digits and writing answers to problems) to predict composite frequencies from measured component frequencies in students. For example, they estimated that writing digits normally occurs at around 1.5 to 2.0 times the frequency of writing answers to basic math problems (Gaasholt, 1970; Haughton, 1972). Thus, if a student writes digits at only 60 per minute, 30 to 40 per minute is probably the range within which that student will be able to write answers to problems. Ratios provide a simple means of quantifying and predicting relations between behavior components and composites.

Predicting special educational needs. Kunzelmann and his associates used frequency (performance measures) and celeration (learning rates) to predict future learning and performance. In the Seattle-Spokane-Tacoma Project (Child Service Demonstration Program, 1974), they collected approximately 150,000 samples of academic behavior frequencies on nearly 3,000 skills from a total of 17,996 students in three school districts. Teachers collected 7 to 10 days of repeated measures per student per skill to determine median frequencies and celerations. By flagging students with less than half the median class frequency or median class celeration on a given skill, it was possible to identify more than 70% of the students, later diagnosed with more costly procedures, as having learning problems. Koenig and Kunzelmann (1980, pp. 49–55) later demonstrated that celeration (learning rate on repeat-
ED measures) is a culturally unbiased predictor of academic success. Several textbooks on assessment (Howell & Kaplan, 1979; White & Haring, 1976) documented basic skills assessment procedures that used both frequency and celeration measures to place students in a curriculum.

Kunzelmann and his associates (Kunzelmann & Koenig, 1980; Magliocea, Rinaldi, Crew, & Kunzelmann, 1977) used single frequency samples of four skills (writing loops, touching circles, touching body parts, counting from 1 to 10) to assess preschool children's readiness for first grade. With 92% predictive validity, children who performed in the bottom 25% in any three of the four skills were diagnosed 1 year later as requiring special education programs.

Identifying individuals' best learning channels. Koenig and Kunzelmann (1980) used celeration to assess learning potential in different learning channels in what they termed learning screening, operationalizing what many in the mainstream educational literature have called learning styles. For example, teachers collected repeated 7 to 10 daily measures of writing answers to written arithmetic problems (see/write), saying answers to written problems (hear/write), and saying answers to spoken problems (hear/say). They then used celerations obtained from these repeated measures to predict the learning channels in which individuals would accelerate performance most rapidly.

Component–composite diagnosis. Bourie and Binder (1980; Binder, 1980) applied frequency and celeration assessment methods to the severely disabled, collecting 10 daily frequency samples of 47 component and composite academic and vocational skills in a population of institutionalized students. They used frequency and celeration measures to identify learning and performance deficits. Influenced by Gilbert's (1978) notion of potential for improving performance as the ratio between the best measured performance and the average or typical performance in a population, they coined the term deficit ratio to refer to the ratio dividing competent adult performance levels or REAPS by each individual student's frequency on a given skill. They identified skills with the greatest deficit ratios as being most in need of remediation and most likely to impose ceilings on other skills, prioritizing them for classroom interventions.

Materials development. In addition to developments in assessment methodology, some of these efforts, notably the Seattle-Spokane-Tacoma project (Child Service Demonstration Program, 1974), resulted in thousands of standard practice/assessment sheets for fine slices of academic skills over multiple curriculum areas. These materials subsequently served as an invaluable resource for other precision teaching work around North America, including massive curriculum development efforts in Hastings County, Ontario during the 1970s and the Sacajawea Precision Teaching Project in Great Falls, Montana (Beck, 1979). These early precision teaching materials, including revised versions, are currently available for purchase through Sopris West, a publishing and training company in Longmont, Colorado.

CUMULATIVE DEFICIT
AND GENERATIVITY

The Problem of Cumulative Dysfluency

In the process of collecting data and using frequency aims, precision teachers began to highlight the phenomenon of cumulative dysfluency. A new understanding of educational failure derived from the recognition that behavior components with frequency deficits, despite their accuracy, accumulate when they are layered on top of one another in a curriculum sequence. This accumulation of dysfluent skills limits and may even prevent acquisition of composites that depend on them.

In mathematics, for example, many readers of this article probably experi-
enced increasing difficulty somewhere in the curriculum. Whether this occurred in long division, algebra, or calculus, the point is the same: Cumulative dysfluencies in prerequisite and component skills mounted to make progress through the curriculum increasingly difficult. If performing simple mental arithmetic calculations requires more than a fraction of a second, for example, then one is likely to experience difficulty when attempting to follow a teacher’s rapid demonstration of solving an algebra equation. Fluency-oriented educators are coming to view cumulative dysfluency as perhaps the single most important factor in long-term student failure (Binder, 1988b; Johnson & Layng, 1992; Pannypacker & Binder, 1992). The analysis of cumulative dysfluency is rich with opportunities for controlled research and for communication with the larger educational community.

Unfortunately, in an educational environment in which accuracy is the only metric for mastery, it is impossible to detect dysfluency prior to its ultimate cumulative effect: the inability to learn or perform complex skills due to multiple dysfluent (but possibly accurate) prerequisites or components. Although measures of performance frequency clearly separate individuals or groups with obviously different levels of competence, accuracy assessments may not (Barrett, 1979).

Generativity: A Result of Cumulative Fluency

After the original discovery that component skills must be fluent to support easy application (Haughton, 1972a), precision teachers began to understand that the ability to combine response classes and improvise or problem solve depends on the development of fluent components.Accounts ofchild development based on an analysis of component behavior frequencies in the womb (Edwards & Edwards, 1970) and in early childhood (Mira, 1977) reflected a view that composite milestones in the traditional account of child development emerge when behavior components increase in frequency and spontaneously combine and are then reinforced by natural consequences. Early investigation of the effects of frequency building on application and transfer of skills (Binder, 1976) emphasized the implications for “creativity” of building high frequencies of behavior components. Classroom programs (Binder, 1979c; Pollard, 1979; Solsten & McManus, 1979) in which students were sometimes able to emit behavior composites without direct training after building frequencies of components, demonstrated this “creative” potential in component-composite relations. Haughton’s (1979, personal communication) attention to the work of deBono (1970) reflected an interest in identifying components that must be made fluent to support flexible, problem-solving, and creative repertoires. Johnson and Layng (1992; Layng, Jackson, & Robbins, 1992) have recently linked the selectionist language of basic research on contingency, adduction and generativity (Andronis, 1983; Epstein, 1985) to the fluency-based instructional design, and have made this “generative” effect of building component frequencies the hallmark of their instructional design model (Johnson & Layng, 1994).

Haughton (1982, personal communication) hypothesized that the more types of behavior an individual can perform at high frequencies, the more likely he or she will be able to learn new classes of behavior and adapt to new situations. He proposed a conceptual formula for the og, calculated as the number of response classes in a repertoire multiplied by the frequencies at which they can be performed. (The term was a play on Ogden Lindsley’s name and by analogy to the erg, a measure of work in the metric system.) He suggested that ogs might predict learning ability for an individual, and that a key to accelerating learning ability is to maximize the frequency of as many critical behavior components as possi-
be in the repertoire of an individual. Whether or not such a calculation is practically feasible, the concept corresponds to the principle of generativity based on frequencies of behavior components.

**AFFECTIVE CORRELATES OF FLUENCY**

The affective correlates of fluent behavior have been a topic of increasing interest among fluency researchers. Binder et al. (1990) observed that students may emit inappropriate or aggressive behavior when asked to perform dysfluent skills for more than brief periods, and that such behavior disappears when teachers shorten practice durations. Binder (1990a), discussing the affective correlates of fluency, observed that salespeople report feeling more confident after achieving fluent verbal behavior required for their jobs. Lindsley (1992) points out that fluency is fun. These observations are consistent with Haughton’s (1980) guideline that above a certain performance level, estimated as about half of REAPS, individuals will continue to practice with little or no explicit feedback or arranged reinforcing consequences other than the usual chart-based feedback procedures. The affective correlates of fluency deserve further research, and may provide a basis for analyzing some of the concepts related to “inner motivation” espoused by traditional educators.

**REMOVING CEILINGS: TOWARD A FLUENCY-BUILDING TECHNOLOGY**

Work at B. H. Barrett’s Behavior Prosthesis Laboratory in Waltham, Massachusetts, during the 1970s and early 1980s focused on development of frequency-based instructional technology. Early laboratory studies (Barrett, 1965, 1969, 1971; Barrett & Lindsley, 1962) and subsequent classroom application (Barrett, 1977b) set the stage for this work by emphasizing the free aspect of free-operant conditioning and the necessity of producing minimal behavior frequencies in order for natural reinforcement contingencies to be effective. Our research over the course of nearly a decade followed a progressive investigation of four kinds of ceilings that can prevent or inhibit development of skill (Binder, 1978b) and methods for removing them. Although this section covers some of the same ground as previous sections, it is useful because framing that work in the context of the four ceilings explains the conceptual and technological evolution in a way that can be applied progressively to improve any type of learning program.

**Removing Measurement-Defined Ceilings**

Most instructional procedures at the time used controlled operant trials, with an emphasis on stimulus control and an accuracy-only approach to measurement. Following programmed instruction, these procedures did not measure behavior frequencies, because they were mostly controlled by the experimenter or teacher. In contrast, we always used frequency measures, even in conjunction with teacher-controlled trials procedures. We compared the actual frequencies of performance allowed by procedures (trials per minute) with the frequencies of similar skills that might occur as free operants in natural settings (responses per minute). For example, typical trials procedures for teaching reading to severely disabled students occurred at 12 or fewer opportunities to respond per minute (Binder, 1977, 1978b), measured with an uninterrupted timer. By comparison, competent oral readers perform at 250 or more words per minute. It became obvious that accuracy-based measurement procedures are not sensitive to these important differences in behavior and instructional procedures (Barrett, 1979). In that sense, they impose a ceiling on educators’ ability to detect the difference between competent and incompetent performance, and may be
among the most basic obstructions to effective educational practice in schools and training programs. Frequency measures allowed us to break through ceilings imposed by the 100% correct maximum in traditional educational assessment methods and to achieve a new level of measurement sensitivity.

Eliminating Procedure-Imposed Ceilings

The identification and removal of measurement-defined ceilings made procedure-imposed ceilings obvious, and led to an effort over the course of several years to design teaching and practice procedures that freed students to behave at their own pace. Much of the instructional time in trials procedures involves a teacher or experimental apparatus presenting stimuli one trial at a time, presenting consequences, and recording responses. Students are repeatedly required to stop and start behaving in a way that does not mimic the normal stream of behavior. Working with prevocational and preacademic skills, we created materials and procedures designed to allow free-operant responding. For example, by shifting from a trials procedure for teaching word naming one word at a time to a procedure in which words were laid out in an array for students to name as rapidly as possible, we were able to instantly triple behavior frequencies in some students without any additional intervention (Binder, 1979d; George, 1975; Pease & George, 1975). Other examples included procedures that allowed students to continuously count objects into arrays of cups marked with numbers, write numbers and letters on practice sheets, and practice components of dressing skills by inserting the arm into multiple cutaway shirt sleeves, one on top of another, as rapidly as possible. With these procedures it was possible to fade out extrinsic antecedent prompts, corrective feedback, and other trial-by-trial interruptions as rapidly as possible, and develop procedures to accelerate uninterrupted correct responding. We used response opportunities per minute as an indicator of instructional efficiency.

Our goals were threefold: (a) to free the behaver to respond at his or her own highest frequency, (b) to provide many opportunities to respond, and (c) to eliminate interruptions of attention involved with trials procedures. These transitions from trials to self-paced procedures often occurred after students had achieved between 66% and 100% accuracy, although we were able to change some acquisition procedures even before achieving accuracy.

Beyond the immediate multiplication of frequency allowed by changes in procedure, it became possible to use both antecedent stimuli and consequences to build frequencies of correct responding without slowing down behavior. High-paced prompting with finger-pointing, touching, and verbal cues often accelerated behavior frequencies. Accentuating amount of work completed (e.g., markers every nth item on an array of stimuli) and amount of time passed (e.g., large sweep-second hands on darkroom timers) enhanced the effects on behavior frequencies of fixed-ratio reinforcement schedules. Procedures called coaching and cheerleading combined energetic "hustling" antecedents with enthusiastic social consequences, often making the classroom appear more like an athletic gym than a school (Binder, 1976, 1977; Binder & Haughton, 1982). The term fluency coaching may have originated with these procedures.

During the same period we observed that performance durations as brief as 5 or 10 min were too long for students who had not achieved minimal behavior frequencies (Binder, 1982, 1984; Binder et al., 1990). Shortening practice durations accelerated and stabilized performance in many cases, often reducing error frequencies without any other intervention.

An unexpected side-effect of these procedures was a significant reduction in problem behavior. Students who had exhibited such inappropriate behavior
as jumping up from their seats, biting their hands, or throwing objects often stopped exhibiting these types of behavior when allowed and encouraged to behave continuously without interruption for brief intervals. Positive affect in the form of smiles and laughter often replaced negative behavior.

These developments would probably not have occurred had we been working with nondisabled students. Most precision teaching procedures in regular public school classrooms already involved relatively self-paced behavior using such materials as practice sheets or pages of reading material. Only with the severely disabled, for whom trials procedures were the norm, did it become obvious how these procedures imposed severe restrictions on the development of competence (Barrett, 1979). Having recognized these restrictions, however, we became more sensitive to analogous limitations imposed by materials and procedures in regular classrooms. For example, competent adults can write as many as 120 answers to simple math problems in a minute, but most public school classrooms neither allow nor encourage students to complete that much work in a single practice episode. Even most classroom procedures for professional adults prevent individuals from responding at optimal pace. It is an essential feature of fluency-based instruction to remove such procedure-imposed ceilings as rapidly as possible.

Remediating Deficit-Imposed Ceilings

Only with procedures that allowed behavior frequencies to seek their own levels did it become clear that students were unable to perform certain behavior components at competent frequencies. Resnick, Wang, and Kaplan (1973) influenced our work by providing an example of thorough component-composite task analysis in basic mathematics, which we adapted for use with severely disabled students (Pease, 1975). The focus of research shifted to assessment and remediation of component behavior deficits and the effects on composite behavior of increasing component frequencies. We worked closely with Haughton and his associates and with colleagues in the Boston area (Binder, 1978a, 1979a; Binder & Haughton, 1982; Bourie & Binder, 1980; Pollard, 1979; Solsten & McManus, 1979). By introducing precision teaching methods to occupational therapists (Binder, 1981b), language therapists (Binder, 1981a; Burgoyne, 1982), physical therapists (Imbriglio, 1992), and entire multidisciplinary teams (Binder & Pollard, 1982; Pollard & Binder, 1983), we established a common measurement language and assessment strategy for identifying and remediating deficit ratios between students' behavior frequencies and aims based on competent adult performance.

We created free-operant analogues of mediated transfer stimulus equivalence procedures (Binder, 1979d) and engaged in research using free-operant bursts on computer keyboards to investigate the effects on chains of increasing the frequencies of smaller keystroke sequences (Blakeslee, Barrett, & Buchman, 1985). This work expanded component-composite analysis beyond academic skills to broader repertoires of complex behavior and focused on remediation of component behavior deficits.

Handicap-Defined Ceilings

Finally, having removed the first three types of ceilings, we acknowledged the existence of nonremediable component dysfluencies in disabled repertoires. Although accuracy measures were often incapable of distinguishing between obviously disabled behavior and competent performance, Barrett (1979) demonstrated that frequency measures of freely emitted behavior could help to define success or failure in application of the then-popular "normalization principle" in special education. In the face of persistent dysfluencies, the alternative was to identify alternative repertoires or to
create "behavior prostheses" to compensate for the deficits (Barrett, 1977a).

**Fluency Blockers and Builders in Instructional Design**

Some time after active fluency research had ceased at the Behavior Prosthesis Laboratory, Binder (1989, 1990a) classified fluency blockers and fluency builders according to categories originating in the ceilings defined in the laboratory (Figure 6). These were factors in the instructional or performance environment related to measurement, procedures, materials, skills (component responses), and knowledge (component discriminations and verbal behavior). As a framework for diagnosing problems limiting any type of instructional program, these categories serve as a useful checklist for improving instructional effectiveness.

In retrospect, an important contribution of our research at the Behavior Prosthesis Laboratory was the introduction of fluency concepts and methods to Kent Johnson during nearly 2 years of informal discussion toward the end of the 1970s. Johnson and his associates (Johnson & Layng, 1992, 1994) have subsequently made important contributions by integrating and expanding many of these concepts and methods in the Morningside model of generative instruction.

**CORROBORATING EVIDENCE OUTSIDE BEHAVIOR ANALYSIS**

Evidence in other traditions of research and educational practice supports many of the general conclusions suggested in this article. Robbins (1994) recently conducted a review of the cognitive literature on automaticity (Bloom, 1986), and previous reference to LaBerge and Samuels (1974) in this article acknowledges the impact of that work on our study of endurance and distractibility. Binder (1979b, 1979c) conducted an extensive review, available on request, that aimed to uncover all precedents for using time-based measures in research on human learning outside the operant conditioning and precision teaching literatures. Although a complete recitation of that review exceeds the scope of this article, no account of the evolution of fluency concepts and methods would be complete without some reference to precedents and influences from that literature.

**Overlearning Trials Procedures and Latency Measures**

Traditional verbal learning and perceptual-motor learning researchers have used the term overlearning to refer to procedures that provide learning trials beyond the point at which learners achieve 100% accuracy (Fitts, 1964; Hall, 1971; Kruger, 1929). The inherent problem with these procedures is that with accuracy-only measures it is impossible to directly assess the effects of overlearning trials beyond the point of 100% accuracy. Instead, experimenters used secondary effects of these trials, such as improvements in retention or transfer of training, as indirect indicators of learning beyond the 100% correct ceiling (e.g., Hall & Wenderoth, 1972; Kruger, 1929).

Comparatively few early researchers measured latencies in paired associate or other forms of verbal learning, probably because the instrumentation required was expensive and inaccurate until recent years, and because subjects require extensive pretrained to respond with reliable speed in trials procedures (Runquist, 1966). However, those who did measure latencies consistently found that they continued to decline with learning trials beyond 100% accuracy and that shorter latencies predict better retention and transfer of training (Hall & Wenderoth, 1972; Judd & Glaser, 1967, 1969; Keller, Thomson, & Tweedy, 1967; Milledge, 1964; Osgood, 1946; Peterson, 1965; Suppes, Groen, & Schlag-Ray, 1966; Theios, 1973). Osgood (1946), for example, concluded that latency
<table>
<thead>
<tr>
<th>Category</th>
<th>Fluency Blockers</th>
<th>Fluency Builders</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measurement</td>
<td>Measurement procedures that ignore the time dimension.</td>
<td>Time-based performance measurement and evaluation.</td>
</tr>
<tr>
<td></td>
<td>Measurement procedures with too few response opportunities for the allotted time.</td>
<td>More response opportunities than an expert can complete in the time allowed.</td>
</tr>
<tr>
<td>Procedures</td>
<td>Too few practice opportunities.</td>
<td>Sufficient practice to attain fluency.</td>
</tr>
<tr>
<td></td>
<td>Preventing learners from moving at their own pace.</td>
<td>Self-paced learning and practice procedures.</td>
</tr>
<tr>
<td></td>
<td>Limited response opportunities per minute.</td>
<td>Many opportunities per minute.</td>
</tr>
<tr>
<td></td>
<td>Emphasis on preventing errors during learning.</td>
<td>Treating errors as &quot;learning opportunities.&quot;</td>
</tr>
<tr>
<td>Materials</td>
<td>Too few examples.</td>
<td>Many examples.</td>
</tr>
<tr>
<td></td>
<td>Materials that are difficult to use, waste paper, movement, etc.</td>
<td>Easy-to-manipulate or use, efficient use of paper, space and movement.</td>
</tr>
<tr>
<td></td>
<td>Unnecessarily wordy worksheets and directions.</td>
<td>Succinct worksheets and directions.</td>
</tr>
<tr>
<td></td>
<td>Difficult-to-read and comprehend.</td>
<td>Easy-to-read and comprehend.</td>
</tr>
<tr>
<td>Skill Elements</td>
<td>Critical steps in procedures or chained skills that are not fluent.</td>
<td>Fluent steps in procedures.</td>
</tr>
<tr>
<td></td>
<td>Tool skills or behavior components that are not fluent.</td>
<td>Fluent tool skills or components.</td>
</tr>
<tr>
<td>Knowledge</td>
<td>Prerequisite knowledge that is not &quot;second nature&quot;, or fluent</td>
<td>Fluent prerequisite knowledge (facts, concepts, structures, principles, classifications or processes.)</td>
</tr>
<tr>
<td>Elements</td>
<td>Inability to fluently locate critical information in reference sources.</td>
<td>Ability to use reference systems or job aids fluently, automatically.</td>
</tr>
</tbody>
</table>

© 1992 Carl Binder. All Rights Reserved.

Figure 6. Factors that either prevent or promote fluency, in language intended for corporate instructional designers and performance improvement specialists.

measures provide a "more sensitive indicator of habit strength" (p. 46) than does accuracy-only.

**Measurement Problems in Trials Procedures**

Some traditional researchers identified methodological problems with trials procedures. Hall (1971, p. 429) recognized that practice beyond 100% accuracy represents "a more stringent criterion," yet acknowledged that in the absence of measures beyond the 100% accuracy ceiling, no direct specification of that criterion is possible.

Kruger (1929) tried to scale the value of an overlearning trial by setting the number of trials required for an individual to reach 100% accuracy as a unit, and then providing all individuals
with 1.5 times that number of trials. His hypothesis was that if one over-
learning trial has the same effect as another for a given individual, all sub-
jects given 1.5 times the number of trials required to achieve accuracy would
show approximately the same proportional increases in retention. Results
did not support this hypothesis.

Peterson (1965) argued further that in trials procedures, because latencies
decrease with overlearning trials, the accuracy assigned to a series of trials
depends artifactually on the time allowed to respond after presentation of
the stimulus (the anticipation interval). He wrote “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 a fixed anticipation interval, rela-
tive frequencies of correct response will not be independent of latency” (p.
167). If given a longer anticipation interval, a subject has a higher probabili-
ty of responding correctly than if given a very short interval in which to re-
respond. Thus, accuracy criteria from one experiment to another may not be
comparable, depending on the anticipation intervals. In effect, accuracy-
only assessment does not support a measurement standard, unless other as-
pects of the procedure are carefully controlled. The general conclusion
from these and other studies is that time-based measures are more sensi-
tive and more reliable indicators than is accuracy-only.

Component–Composite Relations
Outside Behavior Analysis

A variety of studies in the perceptual-motor literature corroborate the
finding that increasing performance speed of component behavior produces
improved performance of composites. For example, Bilodeau and Bilodeau
(1954) found that improving speed of performance of a less proficient com-
ponent of a psychomotor task boosts overall composite performance. Gagne
et al. (1950) found positive transfer to

a perceptual motor task from practice
on a discrimination component. Gagne
and Foster (1949) observed positive
transfer from practice on components
of a motor task.

INTEGRATION AND
COMMERCIALIZATION OF
FLUENCY-BASED INSTRUCTION

Much work during the last two decades has been devoted to integration
and commercialization of fluency-based instructional technology. Early
precision teaching demonstration projects persuaded the federal government
to fund dissemination of fluency-based instruction in public schools for nearly
10 years (Beck, 1979; Beck & Clement, 1991). Subsequent work in Ontario,
Florida, and Utah provided more evidence of success.

Private Sector Businesses

During the 1980s and 1990s, a num-
ber of fluency-based instructional tech-
nologists began private sector busi-
nesses aimed at promulgating and sup-
porting further development of these
methods, with the rationale that estab-
lishing a successful commercial enter-
prise offered the best chance of sup-
porting continued research and devel-
ment independent of grants and pub-
lic sector fads (Binder, 1993b; Binder
& Watkins, 1989). These included Mi-
chael Maloney (Quinte Learning Cen-
ters, Ontario), Carl Binder (Precision
Teaching and Management Systems,
Inc., and Product Knowledge Systems,
Inc., Boston), Kent Johnson (Morning-
side Academy, Seattle), Elizabeth
Haughton (Haughton Learning Center,
Napa, California), Ian and Aileen
Spence (Ben Bronz Academy, Hart-
ford, Connecticut), and Anne Desjar-
dins (Cache Valley School, Utah),
among others.

Efforts to Computerize Fluency

A number of efforts to computerize
fluency-based instruction have met
with mixed results. Maloney and Sum-
BEHAVIORAL FLUENCY

Mithers (1982) produced Mighty Math®, one of the first commercially available software packages for developing fluency in basic academic skills. Ben Bronz Academy has also developed computer programs for practicing basic math skills (Spence, 1996). Perhaps the most sophisticated enterprise in this area, BehaviorTech, Inc. (Orgel, 1984), produced a computer-based learning system called Exemplar® that ultimately failed in the corporate marketplace due to lack of interest rather than lack of results. Claudia McDade and her colleagues (including former colleague Charles Olander) at the Center for Individualized Instruction at Jacksonville State University have produced a number of computer-based testing and instructional software packages. Joseph Parsons, of the University of Victoria, developed ThinkFast™; James Cowardin, John Eshleman, and their associates produced a computer-based training system (now owned by Precision Learning Systems, Inc., Atlanta) aimed at producing fluency. The consistent challenge in these and other efforts to build fluency with computers has been to escape the limitations of controlled operant procedures and to raise ceilings on the speed at which learners can interact with computers in a continuous stream of behavior. In addition, most current computer-based fluency programs suffer from ceilings imposed by component typing-skills dysfluencies among most learner populations. High-speed voice-recognition technology may offer hope for overcoming this problem in the future.

CONCLUSION

Little known to most of our behavioral colleagues, there is a rich history of conceptual and technical evolution focused on development of behavioral fluency. This article has attempted to summarize the key stages and aspects of that evolution for the benefit of those only recently becoming interested in this field.

Fluency offers a new organizing framework (or paradigm) for researchers and practitioners accustomed to accuracy-only measures of educational mastery and deficit and accuracy-only criteria for advancing through curriculum sequences. Although a great deal more systematic research and development must take place in order to pin down key variables and parameters, there is no question that the addition of frequency aims and frequency-building procedures can improve instructional efficiency and effectiveness.

Paradoxically, fluency-based instruction represents a return to what nearly all cultures and individuals with traditions of skilled performance already know: Fluent, well-practiced behavior is the characteristic of true mastery in any field of skilled endeavor. Musicians, athletes, martial artists, skilled craftspeople, and many others already understand the importance of fluency and practice. For behavior analysts, the challenge is to incorporate these principles into our research agendas and technologies, augmenting the methodology of free-operant conditioning with a fresh understanding of component-composite behavior relations.

It may not be too optimistic to predict that with continued and accelerated development of fluency-based instructional technology, many of our most pressing educational problems will become far less daunting. Such a development, however, depends on broad cultural appreciation of this new understanding, a goal now being addressed by some of those involved in fluency research and practice (Binder, 1993b; Binder & Watkins, 1989; Pennypacker & Binder, 1992).

Fluency is a new paradigm for research to the extent that it integrates and redirects our scientific and technological endeavors with a new definition of mastery—one that requires inclusion of the time dimension. It is a new paradigm in education to the extent that it changes teaching practices and enables us to multiply the cost ef-
fectiveness of education and training programs.

REFERENCES


Binder, C. (1978b, October). Four kinds of ceilings which constrain the growth of skill. Presented at a meeting of the American Association for the Education of the Severely and Profoundly Handicapped, Baltimore, MD.


Binder, C. (1979d, June). Response rate measurement in a mediated transfer paradigm: teaching severely retarded students to read. Presented at a meeting of the Association for Behavior Analysis, Dearborn, MI.


Binder, C. (1990b). Precision teaching and cur-
BEHAVIORAL FLUENCY 193

ros-
ter-

vill-

and-

and-

ive-

far-

and-

it.

that-

ing-

ing-

ch-

ance,

That-

ning-

per-

Un-

Fa-

aan-

sion-

Inc.

ncy-

per-

Na-

ruc-

Academic

ork-

MA:

ems,

the

cur-


Binder, C. (1994). To build fluency: practice don’t role play. Presented at a meeting of the National Council on Performance and Instruction, San Francisco. (Handout available from PT/IMS, Inc., P.O. Box 95009, Nonantum, MA 02195)


Haughton, E. C. (1972b). Define your act and set your fluency goals in personal, social, and academic areas. Special Education in Canada, 48(2), 10–11.


Kovacs, M. (Ed.). (1978). *Children of Aggression: A practical taxonomy of normal body control and beginning skills*. Ontario, Canada: Hastings County Board of Education. (Available from M. Kovacs, RR 15 Rd, St. Ann's, Ontario, Canada LOR 1Y0)


ional development and behavior analysis program. Annual report (pp. 52–55). Belmont, MA: Behavior Prosthesis Laboratory, Fernald State School.


Pollard, J. (1979, October). A task analytic approach to teaching the profoundly retarded effective self-initiated toilet paper skills. Presented at a meeting of the American Association for the Education of the Severely and Profoundly Handicapped, Chicago, IL.


Robbins, J. K. (1994, May). Fluency and automaticity: Samenesses and distinctions. Presented at the annual meeting of the Association for Behavior Analysis, Atlanta, GA.


sley, O.

ermen-

tical Dis-

zional

AME-

patria grant

dic Law

us, 871

1)

1 of the

ory and

y MA:

ies.

r about

in J. B.

y doing

9). Ar-

Chil-

toward

(Ed.),

p. 389-

cations.

ing and

. Iowa

riculum

ed com-

ue Cur-

62, Eu-

s to de-

(Avail-

limited, 33)

s to de-

itable

1, Inc.,

s to de-

from O. Box

ay, M.

acy in

Math-

ements

ory and

ology of

York:

1). Ana-

instruc-

pies.

vists.

a study

rate of

on facts

ation].

tic ap-


Wood, S., Burke, L., Kunzelmann, H., & Ko-

enig, C. (1978). Functional criteria in basic math proficiency. Journal of Special Educa-

Technology, 2, 29–36.