

A Nonlinear Approach to Curriculum Design The Role of Behavior Analysis in Building an Effective Reading Program

Janet S. Twyman
Greg Stikeleather

T. V. Joe Layng
Kelly A. Hobbins

Whether teaching reading is rocket science may be debated, but clearly, teaching it effectively, efficiently, and across a variety of learners *is* learning science. In this congressionally recognized “Decade of Behavior” (see www.APA.org), it is becoming increasingly clear that if U.S. education is going to improve, we will have to begin applying principles and procedures derived from the learning sciences directly to teaching. This trend toward recognizing the critical importance of a scientifically informed approach to teaching is indicated by the new federal requirements for scientifically based instruction and in positions such as those represented in the U.S. Department of Education’s (1993) publication, *Toward a New Science of Instruction: Programmatic Investigations in Cognitive Science and Education*.

Headsprout is an Internet-based learning sciences company whose initial product is Headsprout Early Reading. The program is composed of two parts, Headsprout Reading Basics and Headsprout Reading Independence. Whereas Reading Basics focuses on decoding and begin-

ning comprehension and is ideal for Pre-K as well as beginning K-1 learners, Reading Independence has a focus on reading fluency and comprehension. Headsprout has not only applied the principles derived from the scientific study of learning to the teaching of fundamental reading skills, but has also turned the process of building such a program into a science itself, with valid, empirical, replicable results. With over 40% of the nation’s fourth-grade students failing to demonstrate basic reading skills (National Center for Education Statistics, 2004), the need for such a program has become increasingly clear in the past few years (see Layng, Twyman, & Stikeleather, in press).

But what do we mean when we talk about turning the process of building such a program into a science itself? For that matter, what do we mean by *science*? This chapter will describe our approach to science and its application to the development, not just the testing, of instructional programs.

Science is not simply a collection of facts, observations, interpretations, or theories (Brown-

owski, 1956). Science is the *process* by which we develop and test our interpretations or theories. Science is a process of searching for fundamental and universal principles that can offer parsimonious accounts of how observed changes in one set of conditions can result in changes in another set of conditions. The process involves building, testing, and connecting *falsifiable accounts* to describe, explain, and often predict phenomena (Brownowski, 1956; Kitchner, 1982; Sidman, 1960). The scientific method often includes inferences, repeatable experiments, and observations that select one set of inferences over others, and the positing of new inferences (Platt, 1964). The prime criterion in determining the usefulness of an inference is the extent to which the inference correctly makes predictions or explains phenomena verifiable by independent observers and stands up to tests of falsifiability (after Goldiamond & Thompson, 1967 [revision by Andronis, in press]; Popper, 1968). Specifically:

Inferences must be falsifiable. There must be a way to prove the inference wrong. If we can't prove it wrong, it is not a scientific theory. This idea of an inference being falsifiable is one of the most important aspects of science.

Inferences must be able to predict. All scientific inferences must have some predictive value or generality. Prediction can come from the ability to control relevant variables or from the likelihood of a match between predicted and obtained observations. The greater the extent to which predictions hold under a variety of conditions, the greater the generality.

Inferences must be economical. An inference must provide an account of the observed relations that is as complete as possible with the simplest set of principles or axioms possible. In other words, the inference must be parsimonious.

Inferences must be replicable. It is not acceptable that only one person, or only one group, can obtain results that support the inference.

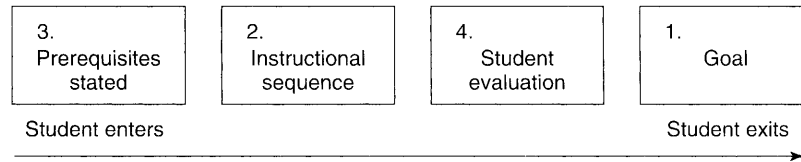
Anyone using proper procedures must be able to achieve the same results.

Inferences must engender confidence. We have degrees of confidence in inferences. Scientists have confidence in inferences that remain intact after repeated attempts at falsification and fit with other inferences and observations (Kitchner, 1982). Although strong confidence in an inference is sometimes warranted, it can never be absolute.

The scientific method can be applied to designing good instruction. *Inferences* about what and how to teach are developed from the current knowledge base. Instructional procedures are designed, tested, evaluated, revised, and retested until a highly *predictable* result is obtained. *Falsified*, or ineffective, strategies are modified or discarded. Patterns of behavior change are identified and connected, with the most effective and simplest strategies for producing the change retained (*parsimony*). Finally, the effects of the instructional methods are verified across a variety of individual learners in different contexts, even across time and space (*replication* and *confidence*).

Headsprout's inferences about how to build a reading repertoire fit within the scientific process. Headsprout used data from more than 30 million responses, across more than 10,000 learners, in a scientific approach to the design of its early reading program. Through the application of single-subject experimental design (Johnston & Pennypacker, 1993; Neuman & McCormick, 1995, 2002; Sidman, 1960), some of our inferences were falsified, whereas others were repeatedly supported by the data. These data allowed us to predict other behavior-environment relations and to replicate them over the project's large scale. However, the process from where we began to the current design did not follow a straight path. As is often true of laboratory research, our approach was (and still is) nonlinear, circuitous, and replete with continuous evaluation and revision.

Figure 3-1.
Numbers 1–4 indicate the steps followed in typical linear instructional design.



NONLINEAR INSTRUCTIONAL DESIGN

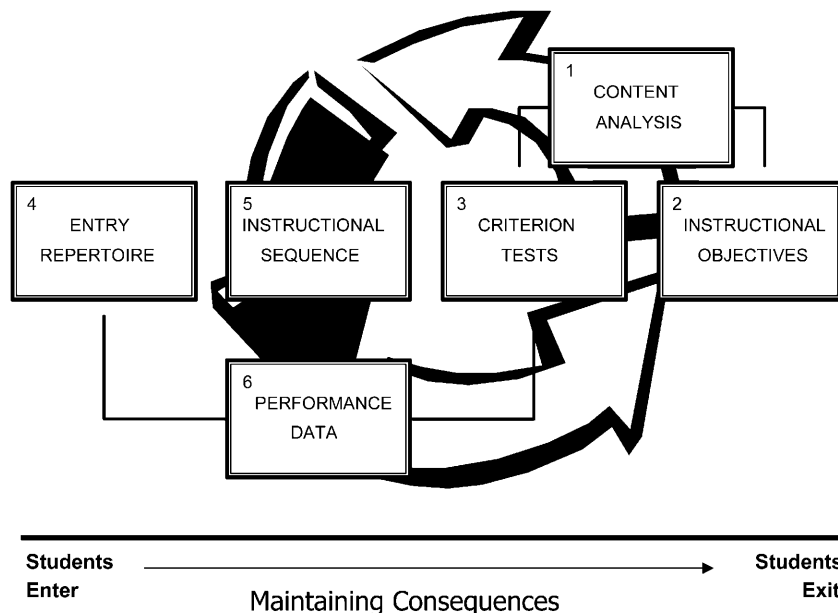
We use the term *nonlinear* in two ways. The first applies to the overall development strategy. The sequence of design and development (see Figures 3-1 and 3-2) did not proceed in a step-by-step, linear manner. The second sense of nonlinear is borrowed from Israel Goldiamond (1975, 1979, 1984) to refer to occasions in which the “behavior of interest” (the target behavior, the terminal behavior, etc.) is not simply a function of the occasion and consequences (and their history) that immediately surround it, but of the occasions and consequences of alterna-

tive behavior patterns (and their history) as well. We will first consider the former meaning.

THE NONLINEAR PROGRAMMING PROCESS

One typical approach to instructional design is to apply a top-down process (see Figure 3-1). The goal is identified, broken down into smaller steps, and checked for “social agreement” (do experts, or at least those with some familiarity, think it’s reasonable?); the program is then written in its entirety and tested with students. At times, designers will take data, note if their

Figure 3-2.
Numbers 1–6 indicate the nonlinear steps adapted from Markle and Tie-mann’s (1967) programming process.



students fail, and redo some portion of the program. This often occurs in the context of field testing with groups of learners. If the overall group tends to meet the goal (often judged by consensus and learner emotional reaction), then the product is considered finished. Many designers and curriculum publishers fail to perform the last two steps of minimal test and revision. Such a program may purport to present content that is derived from scientific principles, but the program itself does not meet the criteria for a scientifically developed program. The program may even provide a better outcome than some alternative approach against which it is compared, but still the program cannot be considered scientifically developed.

Markle and Tiemann (1967) described a different instructional programming process. They noted that the entire instructional design process determines whether an instructional product will fulfill its vision. Markle and Tiemann took the position that a rigorous scientific control-analysis system is necessary for successful instructional design. Nevertheless, that recommendation has seldom been followed (Cook, 1983; Markle, 1969, 1990). One reason may be that there are few examples of its application on a large scale that can serve as a guide to others who are interested in producing quality instructional materials. Another, perhaps greater, obstacle is the time and expertise required to fully implement all elements of a scientific instructional design process. Markle and Tiemann's programming process can be slightly updated and summarized as follows (see Figure 3-2):

1. Perform a content analysis. Content is examined and classified as to the type of learning involved (e.g., strategy, principle, concept, verbal repertoire, sequence or algorithm, multiple discrimination, paired associate, kinesthetic repertoire, chain or motor response).
2. State the objectives. Clear, measurable objectives are developed that reflect the content analysis and the overall goal of the program.

3. Determine the criterion tests. Tests are constructed against a standard that often involves both accuracy and frequency criteria. The tests are developed for each teaching activity or routine within a lesson segment, for each lesson, for blocks of lessons, and finally for the program.
4. Establish the required entry behavior. Given what is to be learned, determine the skills needed to progress through the program. Entry behaviors are the specific prerequisites skills needed for success, not simply prerequisite experiences (such as taking a "prerequisite" course without actually acquiring the behaviors identified in the course).
5. Build the instructional sequence. The content analysis and the criterion tests are used as a guide to produce instruction that will result in learner behavior that meets specified criteria.
6. Use performance data to continually adjust the instructional sequence (5) until it meets the objectives (2).
7. Build in maintaining consequences, an additional step in the process that was added from Goldiamond (1974). Plan for the different types of motivation that will be required, both program extrinsic and intrinsic (Goldiamond, 1974).

The learner begins at 4 (entry repertoire), goes to 5 (instructional sequence), and is evaluated at 3 (criterion tests) to determine if 2 (instructional objectives) has been reached. As is evident, the student does not progress through a sequence in the same way as the program was built. Nor is the program written in its entirety before it is tested. In this approach the learners' behavior shapes the program until nearly all learners meet the specified criteria.

Headsprout has based its development on this nonlinear, scientific instructional design approach. In the next section we describe the progression through this design process, as well as the revision cycles and phases of empirical

testing. We also describe how the “other” meaning of nonlinear (Goldiamond, 1975, 1979, 1984) is critical to the Headsprout approach.

Step 1: Content Analysis

First, we decided what our program would be about: Program design usually begins with at least a general statement of goals. Goals are often accomplishment based and are general statements of what is to be achieved (Mager, 1997). Our goal was to teach reading and decoding skills at the mid-second-grade level in 24 instructional hours.

We studied the existing literature base. We did a thorough analysis of early reading skills—what skills research determined were essential, what skills interfered with or hampered instruction, and what was currently occurring in reading instruction. In the scientific method, we build our inferences from within a coherent framework of the existing knowledge base—from empirical studies and outcomes that have weathered the test of replication. Headsprout founders spent over a year reviewing the literature and determining areas of instruction, levels of instruction, and types of instruction. Using a constructional approach (after Goldiamond, 1974), we asked, What needs to be established, what repertoire needs to be built, where do we begin? In general, we found that research on effective reading instruction focused on these critical skills (as identified by the National Institute of Child Health and Human Development, 2000): phonemic awareness, phonics, fluency, vocabulary development, and text comprehension.

These skills were analyzed into their constituent elements. The types of learning (kinetic, simple cognitive, complex cognitive, etc.) (Tiemann & Markle, 1991), the hierarchy of skills, and the relation of one skill set to another were determined. The final objectives and the criterion tests for each objective were derived from this analysis. A more complete description of these skills and strategies as analyzed by Headsprout can be found in Layng and colleagues (in press).

Step 2: Instructional Objectives

Based on the content analysis and identification of critical areas for reading success, we identified larger “composite” skills that we knew our learners should be taught. For example, we knew that phonics, or the relation between the letters (graphemes) of written language and the sounds of written language (phonemes), involved the ability to rapidly identify print in the presence of sound, or to reliably produce sound in the presence of print. Table 3-1 provides the sample composite skills and types of discriminations we determined students would need to read and decode at the mid-second-grade level.

These skills involved multiple subsets of skills and behavior sequences, which needed to be broken down into their constituent parts. For example, what does it mean to recognize that sounds make words, words make sentences, and sentences make stories? What behaviors comprise that repertoire? What does the learner actually do to demonstrate the skill? To answer those questions, we needed to determine exactly what *component*, or basic, skills made up our *composite* skills. We needed to identify “pinpoints” that would serve as our learning objectives and the “learning channels” (Houghton, 1980) under which our learners would operate. Pinpoints are considered the relevant behaviors that indicate skilled performance, whereas learning channels are the active portion of a behavioral objective. In the instructional design literature, the term *channel* refers to how stimuli are perceived (“input” behavior, which is usually identified by one of the five senses—see, hear, touch, smell, and taste) and how the learner subsequently demonstrates “knowledge” (“output” behavior, or what the learner does).

With Headsprout Early Reading, for example, we examined this composite behavior: “The learner will recognize that there is a one-to-one correspondence between sounds [phonemes] and print [graphemes].” We then broke it down into its component or “tool” skills. What were the minimal behaviors needed to establish a

Table 3-1
Sample Composite Skills

Composite Skill	Type of Learning
Recognize that there is a one-to-one correspondence between sounds (phonemes) and print (graphemes)	Paired associates, multiple discriminations
Recognize that sounds make words, words make sentences, and sentences make stories	Limited range concepts
Identify the beginning, middle, and ending sounds of words	Multiple discriminations, algorithms
Segment words into sounds	Psychomotor chains, multiple discriminations
Use letter-to-sound correspondence to sound out new words	Principle applying, psychomotor chains
Fluently blend sounds to read hundreds of new words	Principle applying, practice
Use a repertoire of basic word families and patterns to read new words	Principle applying, algorithm following, strategies
Read frequently occurring words automatically and confidently	Paired associate, multiple discrimination practice
Read irregularly spelled words and words containing diphthongs, special vowel spellings, and common word endings	Paired associate, multiple discrimination, principle applying
Self-correct reading mistakes	Principle applying
Independently read aloud early stories with at least 90% word recognition accuracy	Verbal repertoire
Recognize and use the cues of punctuation when reading, including commas, periods, question marks, and quotation marks	Principle applying

single letter–sound relationship for a learner? Some of our conclusions are presented in Table 3-2. A live, interactive sample of how a learner would experience this sequence may be found online at www.headsprout.com/aol/teachers.cfm.

Step 3: Criterion Test

Next, we built our criterion-based measures of learning—what we would use to test for mastery, fluency, and application. In one sense, our computer-based Internet delivery made the identification of the response topography easy—we were not using voice recognition, so all we had was a mouse click. A mouse click

worked well for the task of hearing a sound and then clicking on the correct letters representing that sound. We were, however, faced with the challenge of using mouse clicks to indicate the proper blending of sounds, saying words, and reading sentences. We developed strategies reminiscent of psychophysical indicator response methodology (Goldiamond, 1958, 1962). That is, we devised activities whose correct response (a mouse click in this case) required another unrecorded response to occur. For example, after sounding out a word, a learner might be asked to listen to three different pronunciations of a word by three different Headsprout cartoon characters and to click on

Table 3-2
Sample Minimal Behaviors to Establish Letter–Sound Relationships

Pinpoint (objective)	Skill Hierarchy	Learning Channel (input/output)
Listen to pairings of auditory and visual presentations of a phonetic element	Paired associate	[See/(Hear), then Hear/(See)]
Make observing response to element	Paired associate, psychomotor responding	See • Hear/Click element
Select examples of a phonetic element from example/nonexample array	Multiple discrimination	Hear element • See elements/Click element heard

the character that said the sound “like you did.”

The typical method of building a test to determine what a student has learned after an existing instructional sequence is designed or used is not sufficient to provide the moment-to-moment data required to develop a program scientifically. Every response was recorded and any error was evaluated to determine what was responsible for the error. It was never assumed that errors indicated a lack of ability on the part of the learner; instead, they indicated that a change was required in the program. In addition to moment-to-moment response evaluations, tests were built into the sequence in the form of terminal exercises for each particular teaching activity. Furthermore, episodic tests were designed to test the cumulative effect of an entire lesson, and additional tests were designed to determine overall student performance once they finished the entire sequence.

We observed more than 500 children in our on-site laboratory as they engaged with the program. We measured their vocal responding, overall affect, attention to the program, and approach or escape behaviors, as well as empirical data on performance (accuracy of responding, rate, latency, error patterns, etc.). All learners were videotaped, and the tapes were used to evaluate the developing sequences along with the recorded “click” data.

Step 4: Entry Repertoire

We also needed to determine what skills our students needed to bring to the learning interaction for the instruction to be effective. For example, we had to determine the required current relevant repertoire, or entry behavior. Answers were provided through the content analysis and literature reviews, and by further analyzing component skills, probing actual learners, and using tryout and revision cycles. We determined that the language skills of a typical 4-year-old child; the psychomotor skills necessary to move and click a mouse; and a rudimentary concept of “first,” “next,” “last,” and “not” were all that was required. A mouse practice program was developed that included practice on all these elements as they applied to Headsprout Reading. If a child can successfully navigate this program, it is highly likely that the child has the entry skills needed to learn to read with Headsprout.

Step 5: Instructional Sequence

We had determined both what we thought needed teaching and how we thought we would know if students had learned it. We then had to determine how to teach it. We needed an instructional sequence designed to take learners from their entry repertoire to the terminal outcome. We were alternately hampered

and aided by our choice to have all instruction delivered entirely over the Internet. Unlike books, the Internet allows for direct interaction with the learner's behavior. However, unlike face-to-face instruction, no person would be available to mediate and provide contingent feedback. Continuous online delivery, though not completely foreign to us, presented unique challenges. As noted earlier, not only were we unable to rely on the presence of an adult for instruction, but also we wouldn't be able to hear a word our learners said—and yet we were teaching oral reading.

The guiding principles of our design were nonlinear and systemic. We analyzed the immediately or directly visible contingencies, as well as alternative sets. We did this with the primary behaviors of interest, as well as with the matrices of behaviors that made up the repertoire.

Nonlinear Analysis

A nonlinear approach involves analysis of (a) the contingency of which the target behavior is a member (the direct or linear relations); (b) alternative sets, or matrices, of consequential contingencies, of which the target behavior and currently available alternative patterns are members; and (c) the contingencies or relations that can potentiate the matrices (the nonlinear relations). This analysis is applied to gain an understanding of the patterns of observed learner behavior, which occur as a result of the interaction of these matrices.

In a science of instruction, there is no such thing as a student not learning. Errors are treated as the rational outcome of the current program contingencies (i.e., the response requirements on certain occasions and their consequence) and their alternatives, including the learning history of the learner. Consider an example of this approach as it pertains to reading instruction: A learner is faced with the task of, upon hearing the spoken phonetic unit /ip/, selecting a printed letter pair, such as "ip" from a set of alternative letters representing other phonetic units. This is commonly referred to as a condi-

tional discrimination. We might observe that when the learner hears /ip/, the learner selects "ip" on 90% of the occasions; thus we feel quite comfortable that the conditional discrimination is well established. Upon closer examination, however, we might discover that the learner has learned to reject the other stimuli presented and select "ip" by exclusion, thereby getting it correct. Alternatively, he or she might have learned to choose "ip" when it is surrounded by one set of stimuli (a discrimination based on one set of nonexemplars), and select "ip" by exclusion when surrounded by other sets. In this scenario, hearing /ip/ and seeing "ip" may guide only some of the correct responses, whereas other responses considered correct may be occasioned by seeing sets that are something other than "ip." These two different contingencies may combine to produce a 90% correct rate, but may not truly represent that the target discrimination was learned (Sidman 1980, 1992).

Furthermore, when ambiguous conditions are present, a history with the consequences of calling something an example when it is not (a false alarm), or of failing to call something an example when it is (a miss), can be as influential as the consequences of selecting the example (Robbins, Layng, & Karp, 1995). These and other contingencies (see, for example, Blough, 1972) may act together to determine any given response at any given point in time. It is the job of the learning scientists and instructional designers to consider these factors when designing instruction, testing program segments, revising sequences, and interpreting learner data. It is important, therefore, to carefully control and vary how stimuli are presented, their sequence of presentation, the salience of the stimuli, the learner's history of responding to the alternatives, how the response request to the learner is made, and the consequences of responding to all alternatives (see Ray, 1969; Ray & Sidman, 1970).

The analysis can be more complex because the instructional problem presented is not always the problem to be solved. Learner failure in a particular sequence might not be the result

of the design of that particular sequence, but of the design of preceding sequences, their order of presentation, or the amount of practice. Accordingly, a *systemic* analysis (after Goldiamond, 1984) that goes beyond the sequence where the problem is observed (a *topical* analysis) is often required.

Practice is an important part of the instructional sequence as well. Well-designed “guided” practice firms the skills and leads to learner fluency. Guided practice is not simple repetition; it often involves models, confirmations, and corrections. Practice follows the establishment of components so that they will be readily available for recombination with other equally firm components (Johnson & Layng, 1992, 1994), and is often timed. Timing allows for the greatest number of response opportunities in the shortest period of time and can have properties in its own right that assist in the retention and application of skills. The timed practice and the basic instructional sequence are intertwined, one building on another.

Step 6: Performance Data

Each element of the program had to be continuously tested with new users *during* development; it was not sufficient to simply test for outcomes after the program was developed. No matter how well thought out any sequence might be, learners will show where the analysis, design, or both went wrong. Very systematic try-out and revision cycles are required. Data must be used to track changes and evaluate the various versions tested. Fortunately, computer-based instruction enables the collection of data that can be used to adapt the program to the learner’s responses and helps ensure mastery of the material. Data on the percentage of correct responses, albeit helpful, are not sufficient. We also closely examine the frequency of correct versus incorrect responses (after Lindsley, 1997).

In addition, Headsprout’s on-site user test laboratory allowed us to record, observe, or probe for behaviors that the computer could not

detect. If a student could not perform the terminal exercise for an instructional episode or sequence within an episode, the episode or sequence was revised until the student met the exit criteria. Additional testing of outcomes was performed outside the program to further validate that valid criterion measures were, in fact, built into the program. Once the program sequences and episodes were verified, tests were administered to determine if the sequence itself met the criterion. Changes were made in the program until 90% of the learners met the exit criteria. This process has resulted in over 10,000 data-based program revisions to date.

The process involved (and still involves) constant testing, revision and recycle, and retesting. This process is referred to as “formative evaluation” by Scriven (1974), and as “developmental testing” by Markle (1967), and involves the consensus of best practices, experience, and point of view. The design is initially based on previous research, can come from a variety of disciplines, and may be based on elements implemented in the program and not on the program itself. All elements of the program are tested for effectiveness, and if the criteria are not met, alternative strategies are built and tested. The process iterates until all criteria are met, with performance always measured against a set of standards. The sequencing of program steps and their relation to the learner’s behavior is explicitly identified, thereby generating new knowledge—about both the program and the behavior. This process continues and becomes aggregated as the “chunks” of the program units change in size (e.g., for Headsprout, a segment of a lesson, a lesson, groups of lessons, and the program). The research is based on individuals and therefore can be generalized to individuals (Layng et al., in press; Sidman, 1960).

As a result of the developmental testing process, our instructional objectives have been further refined, with skills added and others removed from the sequence. This process was incorporated in three distinct stages of testing (Markle, 1967): developmental testing (conducted

in-house, in the user test laboratory); validation testing (conducted as a beta release of the program, in controlled remote locations); and field testing (conducted via public launch of the program).

These three phases of empirical testing (developmental, validation, and field testing), identified by Markle (1967) and implemented by Headsprout, are briefly summarized within the following guidelines (see also Johnson, Twyman, & Hobbins, 2001).

Developmental Testing

In the developmental testing phase, the goal is to develop a workable instructional program. For Headsprout, this occurred in our on-site user test laboratory. The instructional designers observed and interacted individually with students while the students used the program. Learners progressed through initial drafts of the program and were encouraged to speak out loud their “reasons” for doing what they were doing (after Markle, 1967); they were often questioned about their behavior.

Developmental testing data were collected on a wide variety of general and precise performance and affective measures, both quantitative and qualitative. The cycle is tryout, revision, tryout, revision, and so on, and always begins with the leanest possible design that might achieve the instructional objective (Gilbert, 1962; Markle, 1990). In developmental testing, the learner’s behavior determines what needs to be added to the program. Our motto was that instruction is like salt: It’s easier to put in than take out! Data-based program interventions included software debugging; improving the writing, graphics, layout, or motivational variables; and, most important, improving the instructional variables to achieve program objectives.

Validation Testing

In the validation testing phase, the goal is to obtain a precise description of performance characteristics as they occur when learning with the

entire product. It focuses less on product development and more on the outcomes of systematic replications across learner demographics. Validation testing occurs in various settings and for Headsprout included both schools and the homes of our learners. Ongoing, moment-by-moment (click-by-click) data were collected, as well as learner performance on various criterion-referenced tests. This enabled us to answer questions such as, Who does or doesn’t learn? How much is learned? and In how much time? Data were analyzed to determine the extent and limit of program effectiveness and to make further refinements in the program.

Field Testing

The goal of the field testing phase is to monitor the use of the product “in the field.” The product is compared to other products, used with a new population or setting, or used under different circumstances. Field testing settings vary widely and with Headsprout included thousands of learners. With Headsprout, moment-by-moment (click-by-click) data were collected in real time, and criterion-referenced and norm-referenced outcome measures were used. Data provided more information about the extent and limit of the program’s effectiveness and are still being used to refine the program. As of this writing, randomized control group studies are being conducted to further ascertain the effectiveness of Headsprout’s reading program.

All three testing phases are employed in the development of Headsprout programs. Although very expensive and time consuming, this commitment to user testing is necessary to build instruction that is truly scientifically based (for more information, see Layng et al., in press). The Internet, although presenting challenges, also presents opportunities. Almost every learner response to the program can be collected and analyzed. As the number of users grows, so too does the ongoing field-test

database. When the program is revised, all learners immediately receive the revision.

Step 7: Maintaining Consequences

As with any instructional sequence, we needed contingencies that would maintain the individual learner's behavior throughout the program. Without human delivery, potent consequential contingencies had to be built into every step of the program (Goldiamond, 1974). This is done by (a) ensuring that the learner is successful and makes visible progress in the program; (b) carefully placing extrinsic consequences such as verbal praise, fun sounds, and brief animations; and (c) ensuring that the skills being learned will be useful (the program-intrinsic consequences obtained from reading). While developing this early reading program, we carefully balanced the instructional (and practice) sequences with the type, frequency, and duration of consequences that would maintain or increase the desired behavior.

CONCLUSION

Teaching involves the coordination of a complex set of teacher and student repertoires. A host of complex interlocking consequential contingencies must be carefully analyzed, sequenced, and arranged. Given the importance of effective reading instruction to America's future, we believed it was necessary to base our beginning reading program not only on the most up-to-date scientific literature, but also on state-of-the-art scientific control-analysis methodology in its design. This effort necessitated a nonlinear approach to both overall program design and the analysis of each response, its occasion, and consequence.

How large a scale was the project that produced the initial Headsprout beginning reading program? It took over 25 people, 3 years, and

over \$4 million to produce the initial 40-episode program and its underlying generative learning technology. The next 40 episodes took an additional year. The team included learning scientists, instructional designers, prototype programmers, copyeditors, creative writers, graphic artists, animators, sound engineers, system architects, system administrators, and Web software engineers. The efforts of all of these individuals had to be coordinated and synchronized as part of an iterated product development system—a nonlinear metasystem, informing and being informed by the nonlinear instructional design system. The result was not only a program that effectively teaches beginning reading, but also a unique, patented instructional technology (Layng et al., 2003).

We have demonstrated that it is possible to develop a complex cumulative repertoire over the Internet (see Johnson & Layng, 1992, 1994), in which component behaviors are systematically established and brought together to make larger behavioral units called composite repertoires. These composites can then combine with other components or composites to make yet other composites. Whereas many approaches to computer-based instruction may appear to limit what a learner experiences, our nonlinear approach frees the learner to respond in ways not typically thought possible for computer-based instruction. Indeed, as described in Layng and colleagues (in press), we can actually engineer successful "discovery learning." We can teach our young speakers to be their own listeners who often correct themselves, and we can reliably teach our learners to make correct articulation of sounds and words without the need for expensive, often cumbersome voice recognition hardware and software. We have provided a highly effective and reliable teaching system that can be easily implemented and maintained. There are few limits on the ultimate scalability of the Headsprout beginning reading program. (For more information on how to access the Headsprout curriculum, visit

www.headsprout.com.) We hope that the Headsprout effort will serve as a model for other large-scale projects whose goals are not only the design and deployment of instruction based on scientifically based principles, but also instruction that is itself a product of a scientific approach to design.

REFERENCES

- Blough, D. S. (1972). Recognition by the pigeon of stimuli varying in two dimensions. *Journal of the Experimental Analysis of Behavior*, *18*, 345–367.
- Brownowski, J. (1956). *Science and human values*. New York: Harper and Row.
- Cook, D. A. (1983). CBT's feet of clay: Questioning the informational transmission model. *Data Training*, *3* (12), 12–17.
- Gilbert, T. F. (1962). Mathetics: The technology of education. *Journal of Mathetics*, *1*, 70–73.
- Goldiamond, I. (1958). Indicators of perception: I. Subliminal perception, subception, unconscious perception: An analysis in terms of psychophysical indicator methodology. *Psychological Bulletin*, *55*, 373–411.
- Goldiamond, I. (1962). Perception. In A. Bachrach (Ed.), *Experimental foundations of clinical psychology* (pp. 28–340). New York: Basic Books.
- Goldiamond, I. (1974). Toward a constructional approach to social problems: Ethical and constitutional issues raised by applied behavior analysis. *Behaviorism*, *2*, 1–84.
- Goldiamond, I. (1975). Alternative sets as a framework for behavioral formulations and research. *Behaviorism*, *3*, 49–86.
- Goldiamond, I. (1979). Behavioral approaches and liaison psychiatry. *Psychiatric Clinics of North America*, *2*, 379–401.
- Goldiamond, I. (1984). Training parent trainers and ethicists in nonlinear analysis of behavior. In R. F. Dangle & A. Polster (Eds.), *Foundations of research and practice* (pp. 504–545). New York: Guilford Press.
- Goldiamond, I., & Thompson, D. (1967 [revision edited by P. T. Andronis, in press]). *The functional analysis of behavior*. Cambridge, MA: Cambridge Center for Behavioral Studies.
- Haughton, E. C. (1980). Practicing practices: Learning by activity. *Journal of Precision Teaching*, *1*, 3–20.
- Johnson, K. R., & Layng, T. V. J. (1992). Breaking the structuralist barrier: Literacy and numeracy with fluency. *American Psychologist*, *47*, 1475–1490.
- Johnson, K. R., & Layng, T. V. J. (1994). The Morning-side model of generative instruction. In R. Gardner III et al. (Eds.), *Behavior analysis in education: Focus on measurably superior instruction* (pp. 173–197). Pacific Grove, CA: Brooks/Cole.
- Johnson, K., Twyman, J. S., & Hobbins, K. (2001, May). Using learner performance to improve a Web-based beginning reading program. Paper presented in G. Bruce (Chair), *Evaluate learning efficiency to design learning efficient programs*, at the 27th annual conference for the Association for Behavior Analysis, New Orleans, LA.
- Johnston, J. M., & Pennypacker, H. S. (1993). *Strategies and tactics of behavioral research* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Kitchner, P. (1982). *Abusing science: The case against creationism*. Cambridge, MA: MIT Press.
- Layng, T. V. J., Johnson, K., Twyman, J. S. Ford, V., Layng, M. P., Gilbert, M., & Stikeleather, G. (2003). *U.S. Patent No. 6,523,007*. Washington, DC: U.S. Patent and Trademark Office.
- Layng, T. V. J., Twyman, J. S., and Stikeleather, G. (in press). Selected for success: How Headsprout Reading Basics teaches beginning reading. In D. J. Moran & R. Malott (Eds.), *Empirically supported educational methods*. St. Louis, MO: Elsevier.
- Lindsley, O. R. (1997). Precise instructional design: Guidelines from Precision Teaching. In C. R. Dills & A. J. Romiszowski (Eds.), *Instructional development paradigms* (pp. 537–554). Upper Saddle River, NJ: Educational Technology Publications.
- Mager, R. F. (1997). *Making instruction work* (2nd ed.). Atlanta, GA: CEP.
- Markle, S. M. (1967). Empirical testing of programs. In P. C. Lange (Ed.), *Programmed instruction: Sixty-sixth yearbook of the National Society for the Study of Education: 2* (pp. 104–138). Chicago: University of Chicago Press.
- Markle, S. M. (1969). *Good frames and bad: A grammar of frame writing* (2nd ed.). New York: Wiley.
- Markle, S. M. (1990). *Designs for instructional designers*. Champaign, IL: Stipes.

- Markle, S. M., & Tiemann, P. W. (1967). Programming is a process [Slide/tape interactive program]. Chicago: University of Illinois.
- National Center for Education Statistics. (2004). Percentage of students, by reading achievement level, grade 4: 1992–2003. Retrieved January 14, 2004, from <http://nces.ed.gov/nationsreportcard/reading/results2003/natachieve-g4.asp>
- National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the Subgroups* (NIH Publication No. 00-4754). Washington, DC: U.S. Government Printing Office.
- Neuman, S. B., & McCormick, S. (1995). *Single-subject experimental research: Applications for literacy*. Newark, DE: International Reading Association.
- Neuman, S. B., & McCormick, S. (2002). A case for single-subject experiments in literacy research. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Methods of literacy research* (pp. 105–118). Mahwah, NJ: Erlbaum.
- Platt, J. R. (1964). Strong inference. *Science*, *146*, 347–353.
- Popper, K. R. (1968). *The logic of scientific discovery*. London: Hutchinson.
- Ray, B. A. (1969). Selective attention: The effects of combining stimuli which control incompatible behavior. *Journal of the Experimental Analysis of Behavior*, *12*, 539–550.
- Ray, B. A., & Sidman, M. (1970). Reinforcement schedules and stimulus control. In W. N. Schoenfeld (Ed.), *The theory of reinforcement schedules* (pp. 187–214). New York: Appleton-Century-Crofts.
- Robbins, J. K., Layng, T. V. J., & Karp, H. (1995). Ambiguity and the abstract tact: A signal detection analysis. *The Analysis of Verbal Behavior*, *12*, 1–11.
- Scriven, M. (1974). Evaluation perspectives and procedures. In J. W. Popham (Ed.), *Evaluation in education: Current application*. Berkeley, CA: McCutchan.
- Sidman, M. (1960). *Tactics of scientific research: Evaluating experimental data in psychology*. Boston: Authors Cooperative.
- Sidman, M. (1980). A note on the measurement of conditional discrimination. *Journal of the Experimental Analysis of Behavior*, *33*, 285–289.
- Sidman, M. (1992). Adventitious control by the location of comparison stimuli in conditional discrimination. *Journal of the Experimental Analysis of Behavior*, *58*, 173–182.
- Tiemann, P. W., & Markle, S. M. (1991). *Analyzing instructional content*. Champaign, IL: Stipes.
- U.S. Department of Education. (1993). *Toward a new science of instruction: Programmatic investigations in cognitive science and education*. <http://www.ed.gov/pubs/InstScience/>

STUDY QUESTIONS AND FOLLOW-UP ACTIVITIES

1. Examine research from a reading curriculum publisher (e.g., Houghton Mifflin, McGraw-Hill). Determine whether the curriculum meets the criteria for a scientifically developed program.
2. Select a composite skill to teach a student. Conduct a component/composite analysis to determine the necessary component skills to teach. Plan a fluency-building lesson in which your student practices the component skills.
3. Create a list of composite skills you would like to teach. Match each composite skill with a type of learning as described in Table 3-1.
4. Select a teaching objective and describe the pinpoint, skill hierarchy, and learning channel. Use Table 3-2 as a guide.
5. Explain to a teacher the advantages of using frequency measures of correct and incorrect student performance compared to percentages of correct performance to guide instruction.
6. Review the research base on one of the following: a published reading curriculum (e.g., Four Blocks, SRA's Direct Instruction), computer-based instruction, or toylike reading instruction tools (e.g., Leapfrog). Compare and contrast these findings with Headsprout's research base.

Study questions and follow-up activities prepared by Charles L. Wood and Susan M. Silvestri.

7. Consider this statement: "It was never assumed that errors indicated a lack of ability on the part of the learner; instead, they indicated that a change was required in the program." What are the implications of this statement for classroom instruction?
8. What are the prerequisite skills for most beginning reading programs? How do they compare to Headsprout's prerequisite language and motor skills?
9. Go to Headsprout's Web site and sample a few lessons. Write a description of the lessons that you could use to explain the structure and benefits of the program to the parent of a child who needs supplemental reading instruction.
10. Stage a mock debate between proponents of scientifically developed instructional programs and program developers who think science is of little value.

Focus on Behavior Analysis in Education

Achievements, Challenges, and Opportunities

Edited by

William L. Heward
Timothy E. Heron
Nancy A. Neef
Stephanie M. Peterson
Diane M. Sainato
Gwendolyn Cartledge
Ralph Gardner III
Lloyd D. Peterson
Susan B. Hersh
The Ohio State University

Jill C. Dardig
Ohio Dominican University

Copyright © 2005 by Pearson Education, Inc., Upper Saddle River, New Jersey 07458. Pearson Prentice Hall. All rights reserved. This publication is protected by Copyright and permission should be obtained from the publisher prior to any prohibited reproduction, storage in a retrieval system, or transmission in any form or by any means, electronic, mechanical, photocopying, recording, or likewise. For information regarding permission(s), write to: Rights and Permissions Department.

Pearson Prentice Hall™ is a trademark of Pearson Education, Inc.

Pearson® is a registered trademark of Pearson plc

Prentice Hall® is a registered trademark of Pearson Education, Inc.

Merrill® is a registered trademark of Pearson Education, Inc.

Pearson Education Ltd.

Pearson Education Singapore, Pte. Ltd.

Pearson Education Canada, Ltd.

Pearson Education—Japan

Pearson Education Australia PTY, Limited

Pearson Education North Asia Ltd.

Pearson Educación de Mexico, S.A. de C.V.

Pearson Education Malaysia, Pte. Ltd.



Upper Saddle River, New Jersey
Columbus, Ohio