Part

III

SOFT TECHNOLOGIES

20

FOUNDATIONS OF PROGRAMMED INSTRUCTION

Barbara Lockee
Virginia Tech

David (Mike) Moore
Virginia Tech

John Burton
Virginia Tech

One can gain appreciable insights to the present day status of the field of instructional technology (IT) from examining its early beginnings and the origins of current practice. Programmed Instruction (PI) was an integral factor in the evolution of the instructional design process, and serves as the foundation for the procedures in which IT professionals now engage for the development of effective learning environments. In fact, the use of the term programming was applied to the production of learning materials long before it was used to describe the design and creation of computerized outputs. Romizowski (1986) states that while PI may not have fulfilled its early promise, "the influence of the Programmed Instruction movement has gone much further and deeper than many in education care to admit" (p. 131). At the very least, PI was the first empirically determined form of instruction and played a prominent role in the convergence of science and education. Equally important is its impact on the evolution of the instructional design and development process.

This chapter addresses the historical origins of PI, its underlying psychological principals and characteristics, and the design process for the generation of programmed materials. Programmed Instruction is renowned as the most investigated form of instruction, leaving behind decades of studies that examine its effectiveness. That history of PI-related inquiry is addressed herein. Finally, the chapter closes with current applications of PI and predictions for its future use.

20.1 HISTORICAL ORIGINS OF PROGRAMMED INSTRUCTION

Probably no single movement has impacted the field of instructional design and technology than Programmed Instruction. It spawned widespread interest, research, and publication; then it was placed as a component within the larger systems movement and, finally, was largely forgotten. In many ways, the arguments and misconceptions of the "golden age" of Programmed Instruction over its conceptual and theoretical underpinnings have had a profound effect on the research and practice of our field—past, present and future. When discussing the underpinnings of Programmed Instruction it is easy to get bogged down in conflicting definitions of what the term means, which leads to disagreements as to when it first began, which leads into the arguments, efficacies and origins of particular concepts, and so forth. Since the work (and personality) of B. F. Skinner is included in the topic, the literature is further complicated by the visual array of misconceptions, misrepresentations, etc. of his work. Suffice it to say that the presentation of our view of the history of PI is just that: our view.

The term, Programmed Instruction, is probably derived from B. F. Skinner's (1954) paper "The Science of Learning and the Art of Teaching" which he presented at the University of Pittsburgh at a conference of *Current Trends in Psychology and the*

Behavioral Sciences. In that presentation, which was published later that same year, Skinner reacted to a 1953 visit to his daughter's fourth-grade arithmetic class (Vargas & Vargas, 1992). Interestingly, this paper written in part from the perspective of an irate parent, without citation or review, became the basis for his controversial (Skinner, 1958) work, "Teaching Machines," and his subsequent (1968a) work, "The Technology of Teaching." In the 1954 work, Skinner listed the problems he saw in the schools using as a specific case "for example, the teaching of arithmetic in the lower grades" (p. 90). In Skinner's view, the teaching of mathematics involves the shaping of many specific verbal behaviors under many sorts of stimulus control, and, "over and above this elaborate repertoire of numerical behavior, most of which is often dismissed as the product of rote learning, the teaching of arithmetic looks forward to those complex serial arrangements involved in original mathematical thinking" (p. 90). In Skinner's view, the schools were unable to accomplish such teaching for four reasons. First, the schools relied on aversive control in the sense that the beyond, "in some rare cases some automatic reinforcement (that) may have resulted from the sheer manipulation of the medium—from the solution of problems on the discovery of the intricacies of the number system "(p. 90) children work to avoid aversive stimulation. As Skinner says, "anyone who visits the lower grades of the average school today will observe that . . . the child . . . is behaving primarily to escape from the threat of . . . the teacher's displeasure, the criticism or ridicule of his classmates, an ignominious showing in a competition, low marks, a trip to the office 'to be talked to' by the principal" (p. 90).

Second, the school did not pay attention to the contingencies of reinforcement; for those students who *did* get answers correct, many minutes to several days may elapse before papers are corrected. He saw this as a particular problem for children in the early stages of learning who depend on the teacher for the reinforcement of being right as opposed to older learners who are able to check their own work.

The third problem that Skinner (1954) noted was "the lack of a skillful *program* which moves forward through a series of progressive approximations to the final complex behavior desired" (p. 91). Such a program would have to provide a lengthy series of contingencies to put the child in possession of the desired mathematical behavior efficiently. Since a teacher does not have time to reinforce each response, he or she must rely on grading blocks of behavior, as on a worksheet. Skinner felt that the responses within such a block should not be related in the sense that one answer depended on another. This made the task of programming education a difficult one.

Finally, Skinner's (1954) "most serious criticism of the current classroom is the relative infrequency of reinforcement" (p. 91). This was inherent in the system since the younger learner was dependent upon the teacher for being correct, and there were a lot of learners per teacher. A single teacher would be able to provide only a few thousand contingencies in the first four years of schooling. Skinner estimated that "efficient mathematical behavior at this level requires something of the order of 25,000 contingencies" (p. 91).

Interestingly, Skinner (1954) felt that the results of the schools' failure in mathematics were not just student incompetence, but anxieties, uncertainties, and apprehensions. Few students ever get to the point where "automatic reinforcements follow as the natural consequence of mathematical behavior. On the contrary, . . . the glimpse of a column of figures, not to say an algebraic symbol or an integral sign, is likely to set off—not mathematical behavior but a reaction of anxiety, guilt or fear" (Skinner, 1954, p. 92). Finally, the weaknesses in educational technologies result in a lowered expectation for skills "in favor of vague achievements—educating for democracy, educating the whole child, educating for life, and so on" (p. 92).

Important to the field of instructional design and technology, Skinner (1954) says "that education is perhaps the most important branch of scientific technology" (p. 93) and that "in the present state of our knowledge of educational practice, scheduling (of behaviors and consequences) appears to be most effectively arranged through the *design* of the material to be learned. He also discusses the potential for mechanical devices to provide more feedback and to free the teacher up from saying right or wrong (marking a set of papers in arithmetic—'Yes, nine and six are fifteen; No, nine and seven are not eighteen—is beneath the dignity of any intelligent individual," (Skinner, 1954, p. 96) in favor of the more important functions of teaching.

In his article "Teaching Machines," published in *Science* (1958a), Skinner pushed harder for the use of technology in education that could present programming material prepared by programmers. This work also discusses the notion that whether good programming is to become a scientific technology, rather than an art, will depend on the use of student performance data to make revisions. Again, he sees the powerful rule that machines could play in collecting these data. Finally, Skinner's (1958a) work has a rather casual, almost *throw off* phrase that generated a great deal of research and controversy:

> In composing material for the machine, the programmer may go directly to the point. A first step is to define the field. A second is to collect technical terms, facts, laws, principles, and cases. These must then be arranged in a plausible developmental order—*linear if possible, branching if necessary* [italics added]. (p. 974)

It may be that Skinner (1954, 1958) was the first to use the vocabulary of programmed materials and designed materials, but it was the rest of his notions which Reiser (2001) says "began what might be called a minor revolution in the field of education" (p. 59) and, according to Heinich (1970) "has been credited by some with introducing the system approach to education" (p. 123). We will briefly examine some of the key concepts.

20.1.1 Teaching Machines

Much of the research regarding Programmed Instruction was based on the use of a *teaching machine* to implement the instructional event. As Benjamin (1988) noted, "the identification of the earliest teaching machine is dependent on one's definition of such machines" (p. 703). According to Benjamin's history, H. Chard filed the first patent for a device to teach reading in 1809. Herbert Akens (a psychologist) patented a device in 1911 that presented material, required a response, and indicated whether the response was right or wrong. The contribution of this device, which was a teaching aid rather than an automatic or self-controlling device, was that it was based on psychological research. In 1914, Maria Montessori filed a patent claim for a device to train the sense of touch (Mellan, 1936, as cited in Casas, 1997). Skinner (1958a) and most others (see, for example, Hartley & Davies, 1978) credit Sidney Pressey. Beginning in the 1920s, Pressey designed machines for administering tests. Hartley and Davies (1978) correctly point out that Pressey's devices were used *after* the instruction took place, but more important to Skinner, however, was Pressey's (1926) understanding that such machines could not only test and score—they could teach. Moreover, Pressey realized that such machines could help teachers who usually know, even in a small classroom, that they are moving too fast for some students and too slow for others.

20.2 PSYCHOLOGICAL PRINCIPLES AND ISSUES

In the limited space available, we will address the primary concepts behind Programmed Instruction and their origins. For reasons of space and clarity, ancillary arguments about whether Socrates or Cicero was the first "programmer," or trying to draw distinctions between reinforcement (presumably artificial) versus feedback (automatic or natural reinforcement) will not be discussed (c.f. Merrill's, 1971, notions on cybernetics, etc.).

Similarly, the issue of overt versus covert responding has been discussed in the chapter on behaviorism in this handbook. Certainly Skinner did not distinguish between private and public behaviors except in terms of the ability of a teacher or social group to deliver consequences for the latter. It is useful to mention the notion of active responding—that is whether the learner should be required to respond at all, and if so, how often. In a behavioral sense, behaving, publicly or privately, is necessary for learning to occur. Some of the discussion may be confounded with the research/discussion on step-size that will be covered later in this chapter (see, e.g., Hartley, 1974). Others were apparently concerned that too much responding could interfere with learning.

Finally, the rather contrived distinction between programmed learning and Programmed Instruction that, for example, Hartley (1974) makes, will not be discussed beyond saying that the presumed target of the argument, Skinner (1963) stated that he was writing about a new pedagogy and the programming of materials grounded in learning theory.

20.2.1 Operational Characteristics of PI

Bullock (1978) describes PI as both a product and a process.

> As a process, PI is used for developing instruction systematically, starting with behavioral objectives and using tryouts of the instruction to make sure that it works satisfactorily As a product, PI has certain key features, such as highly structured sequence of instructional units (frames) with frequent opportunities for the learner to respond via problems, questions, etc. typically accompanied by immediate feedback. (p. 3)

Lysaught and Williams (1963) suggest that Programmed Instruction maintains the following characteristics. First, it is mediated. Beginning as print-based text, Programmed Instruction grew to leverage each new media format as technologies merged and evolved. Also, PI is replicable, as its results consistently produce the same outcomes. It is self-administrating because the learner can engage in the instructional program with little or no assistance. Its self-paced feature allows the learner to work at a rate that is most convenient or appropriate for his or her needs. Also, the learner is required to frequently respond to incrementally presented stimuli, promoting active engagement in the instructional event. PI is designed to provide immediate feedback, informing the learner of the accuracy of his or her response, as well assisting in the identification of challenges at the point of need. Additionally, PI is identified by its structured sequences of instructional units (called frames), designed to control the learner's behavior in responding to the PI.

20.2.2 Linear Versus Branching Systems

The goal of early developers of programmed instruction was to design the instructional activities to minimize the probability of an incorrect response (Beck, 1959). However, much has been made of the distinction between what some have called Crowder's (1960) multiple-choice branching versus Skinner's linear-type program (see, for example, Hartley, 1974). Crowder, like Skinner (1954, 1958a) likens his intrinsic system to a private tutor. Although Crowder himself claimed no theoretical roots, his method of intrinsic programming or "branching," was developed out of his experience as a wartime instructor for the Air Force. Crowder's method used the errors made by the recruits to send them into a different, remedial path or branch of the programming materials. Although the remediation was not in any way based on any sort of analysis of the error patterns or "procedural bugs" (see, for example Brown & VanLehn, 1980; Orey & Burton, 1992) it may well have been the first use of errors in a tutorial system. Although much has been made about the differences between Skinner and Crowder, it is clear that although the two men worked independently, Skinner was clearly aware of the use of branching and accepted it "if necessary" in 1958 (Skinner, 1958a, p. 974). Crowder began publishing his work a year later in 1959 (Crowder, 1959, 1960, 1964). In a sense they were talking about two very different things. Skinner was writing about education and Crowder was writing from his experience in the teaching complex skills to adults with widely varying backgrounds and abilities. The issue is informative, however. Neither man wanted errors per se. Skinner's goal was an error rate not to exceed 5 percent (1954). His intention was to maximize success in part in order to maximize (reinforcement) and, at least as important to minimize the aversive consequences of failure. Crowder (1964) would prefer to minimize errors also, although he accepts an 85 percent success rate (15% error rate). Recalling the context of his learner group that ran at least from college graduates to those with an 8th grade education, Crowder (1964) says:

> Certainly no one would propose to write materials systematically designed to lead the student into errors and anyone would prefer programs in which no student made an error *if this could be achieved*

> *without other undesirable results*. . . . We can produce critically effort-free programs if we are careful never to assume knowledge that the most poorly prepared student does not have, never to give more information per step than the slowest can absorb, and never to require reasoning beyond the capacities of the dullest. The inevitable result of such programs is that the time of the average and better than average is wasted. (p. 149)

In short, Skinner saw errors as a necessary evil—motivational and attention getting, but essentially practicing the wrong behavior and receiving aversive consequences for it. Crowder saw errors as unavoidable given the needs of teaching complex skills to students given different backgrounds and whose ability levels varied form "dull" to "better than average." Crowder's (1960, 1964) contribution was to try to use the errors that students made to try to find the breakdown in learning or the missing prerequisite skill(s).

20.2.3 Objectives

Central to the roots of Programmed Instruction is the idea that programmers must decide what students should be to be able to do once they have completed the program. Generally, this involves some sort of activity analysis and specification of objectives. Dale (1967) traces this approach back to Franklin Bobbitt (1926, as cited in Dale) writings:

> The business of education today is to teach the growing individuals, so far as their original natures will permit, to perform efficiently those activities that constitute the latest and highest level of civilization. Since the latter consists entirely of activities, the objectives of education can be nothing other than activities, and since, after being observed, an activity is mastered by performing it, the process of education must be the observing and performing of activities. (p. 33)

Charters (1924, as cited in Dale, 1967) who, like Bobbitt, was concerned with curriculum and course design contends that objectives are a primary component of the design process. Tyler (1932) used the notions of Charters in his behavioral approach to test construction. Tyler wrote that it was necessary to formulate course objectives in terms of student behavior, establish the situations or contexts in which the students are to indicate the objective, and provide the method of evaluating the student's reactions in light of each objective. Miller (1953, 1962) is generally credited with developing the first detailed task analysis methodology which working with the military (Reiser, 2001). This provided a methodology for taking a complex skill and decomposing it into objectives, sub-objectives, etc. Bloom and his colleagues (Bloom, Englehart, Furst, Hill, & Krathwohl, 1956) created a taxonomy of learner behaviors, and therefore objectives, in the cognitive domain. Robert Gagne (1956) further segmented objectives/behaviors into nine domains. His writings in the area of intellectual skills is consistent with a hierarchy of a taxonomy such that consistent with Skinner (1954, 1958b) subordinate skills need to be mastered in order to proceed to super-ordinate skills. Mager's (1962) work became the bible for writing objectives.

20.2.4 Formative Evaluation

Skinner's (1954, 1958b) early work had indicated the importance of using learner data to make revisions in instructional programs. In a sense, this technology was well established through Tyler's (1932) discussion of the use of objective-based tests to indicate an individual's performance in terms of the unit, lesson, or course objectives (Dale, 1967). Glaser (1965; Glaser & Klaus, 1962) coined the term *criterion-referenced* measurement to differentiate between measures concerned with comparing the individual against a criterion score or specific objectives and *norm-referenced* measurement which ranked the individual's performance compared to other individuals. What was needed, of course, was to change, at least in part, the use of such tests from strictly assessing student performance to evaluating program performance. Indeed, Cambre (1981) states that practitioners such as Lumsdaine, May, and Carpenter were describing methodologies for evaluating instructional materials during the Second World War and beyond. What was left was for Cronbach (1963) to discuss the need for two types of evaluation and for Scriven (1967) to label them formative and summative to distinguish between the efforts during development when the product was still relatively fluid or malleable versus the summative or judgmental testing after development is largely over and the materials are more "set." Markle's (1967) work became a key reference for the formative and summative evaluation of Programmed Instruction.

20.2.5 Learner-Controlled Instruction

Later in the chapter many variations and permutations of Programmed Instruction will be discussed, but one is briefly covered here because it was contemporary with Skinner's and Crowder's work and because it has some special echoes today. Mager's (1962) learner-controlled instruction used the teacher as a resource for answering student questions rather than for presenting material to be learned. Although largely neglected by Mager and others, perhaps in part because the approach or method did not lend itself to objectives (although the students knew them and were held accountable for them) or design, the methodology does resonate with hypermedia development and related research of the last decade. It would be interesting to see if Mager's findings that students prefer, for example, function before structure or concrete before abstract versus instructors who tend to sequence in the other direction.

20.2.6 Transfer of Stimulus Control

At the beginning of the learning sequence, the learner is asked to make responses that are already familiar to him. As the learner proceeds to perform subsequent subject matter activities that build upon but are different from these, learning takes place. In the course of performing these intermediate activities, the student transfers his original responses to new subject matter content and also attaches newly learned responses to new subject matter.

20.2.7 Priming and Prompting

Two terms that were important in the literature and are occasionally confused are priming and prompting. A prime is meant to elicit a behavior that is not likely to occur otherwise so that it may be reinforced. Skinner (1968a) uses imitation as an example of primed behavior. Movement duplication, for example, involves seeing someone do something and then behaving in the same manner. Such behaviors will only be maintained, of course, if they result in reinforcement for the person doing the imitating. Like all behaviors that a teacher reinforces, to be sustained it would have to be naturally reinforced in the environment. Skinner (1968a) also discusses product duplication (such as learning a birdcall or singing a song from the radio) and nonduplicative primes such as verbal instructions. Primes must be eliminated in order for the behavior to be learned.

Prompts are stimulus-context cues that elicit a behavior so that it can be reinforced (in the context of those stimuli). Skinner (1958a) discusses spelling as an example where letters in a word are omitted from various locations and the user required to fill in the missing letter or letters. Like a cloze task in reading, the letters around the missing element serve as prompts. Prompts are faded, or *vanished* (Skinner, 1958a) over time.

20.3 THE DESIGN OF PROGRAMMED INSTRUCTION

While no standardized approach exists for the production of Programmed Instruction (Lange, 1967), some commonalities across approaches can be identified. One author of an early PI development guide even expresses reluctance to define generalized procedures for the creation of such materials, stating that, "there is a dynamic and experimental quality about Programmed Instruction which makes it difficult and possibly undesirable to standardize the procedures except in broad terms" (Green, 1967, p. 61). In fact, the evolution of the instructional design process can be followed in the examination of PI developmental models. Early descriptions of PI procedures began with the selection of materials to be programmed (Green 1967; Lysaught & Williams 1963; Taber 1965). In 1978, long after the establishment of instructional design as a profession, Bullock, (1978) published what Tillman and Glynn (1987) suggest is "perhaps the most readable account of a PI strategy" (p. 43). In this short book, Bullock proposed his ideal approach to the creation of PI materials, the primary difference from earlier authors being the inclusion of a needs assessment phase at the beginning of the process. Additionally, upon the introduction of Crowder's (1960) notion of branching as a programming approach, future authors began to incorporate a decision phase in which programmers had to choose a particular design paradigm to follow—linear, branching, or some variation thereof—before program design could continue (Bullock, 1978; Markle, 1964).

The following description of the program development process incorporates phases and components most common across widely cited models (e.g., Bullock, 1978; Lysaught & Williams, 1963; Markle, 1964; Taber, Glaser, & Schaefer, 1965). However, as mentioned previously, since no standardized model or approach to PI development exists, authors vary on the order and nomenclature in which these steps are presented, so the following phases are offered with the understanding that no standard sequence is intended. (For a graphical examination of the evolution of the PI process, see Hartley, 1974, p. 286.) Early in the program development process, a need for instruction is defined, along with the specification of content and the establishment of terminal performance behaviors or outcomes. Also, characteristics and needs of the target group of learners are analyzed so that the most appropriate starting point and instructional decisions can be made. Following the definition of instructional need and audience, programmers conduct a behavioral analysis to determine the incremental behaviors and tasks that will lead the student to the terminal performance. When more is known about the learners and the instructional need, the program creator selects a programming paradigm, referring to the navigation path in which the learner will engage. Typically the choice is made between linear and branching designs, as previously discussed, however, other variations of these models are described in the following section of this chapter. After the general approach to programming has been decided, the sequencing of content and the construction of programmed sequences, called frames, can begin. Although authors differ on the stage at which evaluation of the initial program should begin (Green, 1967; Lysaught & Williams, 1963; Markle, 1967), feedback is collected from students in trial runs prior to production and program revisions are based on learner feedback. The following section describes each of the aforementioned components of program development.

20.3.1 Specification of Content and Objectives

Most descriptions of the PI development process begin with a determination of what content or topic is to be taught through defining the terminal behavior and, given that, move to the delineation of the program's objectives. Several of the authors' approaches described in this section (Green, 1967; Lysaught & Williams, 1963; Mechner, 1967; Taber et al., 1965) base their discussion of defining terminal behavior and writing effective, measurable objectives on the work of Mager (1962). Once the PI developer clearly specifies the intended outcomes of the program in observable and measurable terms, then the creation of assessment items and evaluation strategies can be planned. Mager's approach to the creation of objectives, through stating what the learner will be able to do as a result of the instruction, the conditions under which the performance can occur, and the extent or level that the performance must be demonstrated, was not only the widely accepted method for PI purposes, but remains the classic approach to objective writing in current instructional design literature.

20.3.2 Learner Analysis

Authors of PI programs sought to collect relevant data about the intended learner group for which the program was to be developed. Such data was related to the learners' intelligence, ability, pre-existing knowledge of the program topic, as well as demographic and motivational information (Lysaught & Williams, 1963). Bullock (1978) describes the target audience analysis as a means to collect information regarding entry-level skills and knowledge to permit design decisions such as pre-requisite content, the program design paradigm, media requirements necessary to support instruction, and selection of representative learners for field tests and program evaluation.

20.3.3 Behavior Analysis

The process of engaging in a behavior analysis for the purpose of sequencing the instruction was commonly advocated in the literature on PI (Mechner, 1967; Taber et al., 1965). Such an analysis served as the early forerunner to the task analysis stage of current instructional design practice. Mechner suggests that most of the behaviors that are usually of interest within education and training can be analyzed in terms of discriminations, generalizations, and chains. Discriminations consist of making distinctions between stimuli. Generalizations address a student's ability to see commonalities or similarities among stimuli. When a learner can make both distinctions and generalizations regarding particular stimuli, that learner is said to have a concept. A chain is a behavioral term for procedure or process. Mechner's definition of chaining is "a sequence of responses where each response creates the stimulus for the next response" (p. 86–87). Once the discriminations, generalizations, and chains are analyzed, the programmer must determine which concepts are essential to include, considering the particular needs, abilities, strengths, and weaknesses of the target audience.

20.3.4 Selection of a Programming Paradigm

Overarching the varied approaches to sequencing PI content is the programmer's decision regarding the linearity of the program. In the early days of PI, heated debates took place over the virtues of linear versus branching programs. Linear, or extrinsic, programs were based on work of B. F. Skinner. Markle (1964) reminds the reader that while a linear design may indicate that a learner works through a program in a straight line, that linear programs also maintain three underlying design attributes—active responding, minimal errors, and knowledge of results.

Lysaught and Williams (1963) present several variations of the linear program that were developed before the notion of branching was developed. Modified linear programs allow for skipping certain sequences when responses have been accurate. Linear programs with sub-linears provide additional sequences of instruction for those who desire extra information for enrichment or supplemental explanation. Linear programs with criterion frames can be used to determine if a student needs to go through a certain sequence of material and can also be used to assign students to certain tracks of instruction.

Intrinsic programming is based on the work of Norman Crowder (1959). "The intrinsic model is designed, through interaction with the student, to present him with adaptive, tutorial instruction based on his previous responses rather than to simply inform him of the correctness or incorrectness of his replies" (Lysaught & Williams, 1963, p. 82). Taber et al. (1965) describe a variation on the intrinsic model, entitled the multitrack program. In a multitrack program, several versions of each frame are designed, each with increasing levels of prompts. If the learner cannot respond accurately to the first frame, s/he is taken to the second level with a stronger prompt. If a correct response still cannot be elicited, the learner is taken to a third level, with an even stronger prompt. This design strategy allows learners who may grasp the concept more quickly to proceed through the program without encountering an unnecessary amount of prompting.

Selection of a paradigm is based on earlier steps in programming process, such as the type of skills, knowledge, or attitudes (SKAs) to be taught, existing assumptions regarding learners, the need for adaptive work, etc. If there is a high variance in ability in a group of learners, then providing options for skipping, criterion frames, or branching would be helpful in supporting individual needs.

20.3.5 Sequencing of Content

Skinner's (1961) article on teaching machines suggested that the one of the ways that the machine helps with teaching is through the orderly presentation of the program, which in turn is required to be constructed in orderly sequences. Following the selection of an overarching programming paradigm, decisions regarding the sequencing of the content can be made. A general PI program sequence is characterized by an introduction, a diagnostic section, an organizing set/theory section (to help learner focus on primary elements of teaching/testing section), a teaching, testing section, practice section, and finally, a review or summary is presented to reinforce all of the concepts addressed in the specific program (Bullock, 1978).

Again, no standard approach exists for the sequencing of content and a variety of models are found in the literature. Lysaught and Williams (1963) describe several techniques, the first of which is the pragmatic approach, or the organization of behavioral objectives into logical sequence. "This order is examined for its internal logic and flow from beginning to end. Often an outline is developed to ensure that all necessary information/steps/components are addressed and that nothing important is omitted" (p. 92).

Another common approach to sequencing content was developed by Evans, Glaser, and Homme (1960), and is known as the RULEG system. The RULEG design is based on assumption that material to be programmed consists of rules or examples. So, the rule is presented, followed by examples and opportunities to practice. In some instances, the reverse approach, EGRUL, is used, presenting the learner with a variety of examples and guiding the behavior to comprehend the rule. Mechner (1967) suggests that the target audience should determine which approach is used. If the concept is simple or straightforward, then learners would likely benefit from the RULEG sequence. If the concept is more abstract or complex, then the EGRUL technique would be the better choice in shaping learner behavior.

In 1960, Barlow created yet another method for PI design in response to his students' dislike for the traditional stimulus-response approach, as they felt the technique was too test-like. Barlow's sequencing method was entitled *conversational chaining,* a reflection of the interconnected nature of the program's frames. The design requires the learner to complete a response to the given stimulus item, but instead of programmatic feedback about the correctness of that response within the stimulus frame; the learner checks his or her accuracy in the following frame. However, the response is not presented separately, but is integrated within the stimulus of the following frame and is typically capitalized so that it is easily identified. As such, the flow of the program is more integrated and capable of eliciting the chain of behavior targeted by the designer.

Another well known, but less widely adopted programming method was developed by Gilbert (1962). This approach, called mathetics, is a more complex implementation of reinforcement theory than other sequencing strategies. This technique is also referred to as *backwards chaining,* since the design is based on beginning with the terminal behavior and working backwards through the process or concept, in step-wise fashion.

20.3.6 Frame Composition

Taber et al. (1965) suggest that a programmed frame could contain the following items: (1) a stimulus which serves to elicit the targeted response, (2) a stimulus context to which the occurrence of a desired response is to be learned, (3) a response which leads the learner to the terminal behavior, and (4) any material necessary to make the frame more readable, understandable, or interesting (p. 90). They also contend that it may not be necessary to include each of these components in every frame. Some frames may contain only information with no opportunity for response, some may be purely directional.

One aspect of the stimulus material that is inherent in Programmed Instruction is the inclusion of a *prompt*. A prompt in Skinner's view (1957) is a supplementary stimulus, which is added to a program (in a frame or step) that makes it easier to answer correctly. The prompt is incapable of producing a "response by itself, but depends upon at least some previous learning" (Markle, 1964, p. 36). Skinner proposes two types of prompts, formal and thematic. Formal prompts are helpful in the introduction of new concepts, as learners may have little or no basis for producing their own, unsupported response. A formal prompt typically provides at least a portion of the targeted response as part of its composition, generating a low-strength response from the learner. Also, the physical arrangement of the frame may serve as a formal prompt type, suggesting to the learner cues about the intended response, such as the number of letters in the response text, underlined words for particular emphasis, the presentation of text to suggest certain patterns, etc. (Taber et al., 1965). Thematic prompts attempt to move the learner toward production and application of the frame's targeted response in more varied contexts in order to strengthen the learner's ability to produce the terminal behavior. Taber et al. describe a variety of design approaches for the creation of thematic prompts. The use of pictures, grammatical structure, synonyms, antonyms, analogies, rules, and examples are all effective strategies that allow the programmer to create instruction that assists the learner in generating the correct response.

The strength of the prompt is another important design consideration and is defined as the likelihood that the learner will be able to produce the targeted response and is influenced by logical and psychological factors related to the design of the frame (Markle, 1964). As new content or concepts are introduced, prompts should be strong to provide enough information so that a correct response can be generated. As low-strength concepts are further developed, prompts can be decreased in strength as learners can rely on newly learned knowledge to produce accurate responses. This reduction and gradual elimination of cues is known as fading or vanishing and is another PI-related phenomenon popularized by Skinner (1958b).

Another design consideration in the programming of frames is the selection of response type. Taber et al. (1965) describe a variety of response type possibilities and factors related to the basis for selecting from constructed answer, multiple choice, true–false, and labeling, to name a few. Also, another response mode option that has been the subject of instructional research is overt versus covert responding. While Skinner (1968a) believes that active responses are necessary and contribute to acquisition of the terminal behavior, others contend that such forced production may make the learning process seem too laborious (Taber et al.). Research addressing this design issue is described in detail later in this chapter.

20.3.7 Evaluation and Revision

As stated earlier, one of the hallmarks of the Programmed Instruction process is its attention to the evaluation and revision of its products. Skinner (1958a) suggested that a specific advantage of Programmed Instruction is the feedback available to the programmer regarding the program's effectiveness; feedback available from the learner through trial runs of the product. In fact, many credit PI with the establishment of the first model of instruction that mandates accountability for learning outcomes (Hartley, 1974; Lange, 1967; Rutkaus, 1987). Reiser (2001) indicates that the PI approach is empirical in nature, as it calls for the collection of data regarding its own effectiveness, therefore allowing for the identification of weaknesses in the program's design and providing the opportunity for revision to improve the quality of the program. Markle (1967) presents perhaps the most explicit procedures for three phases of empirical product evaluation: developmental testing, validation testing, and field-testing. While other authors offer variations on these stages (Lysaught & Williams, 1963; Romiszowski, 1986; Taber et al., 1965), these phases generally represent components of formative and summative evaluation.

What factors should one consider when attempting to determine the effectiveness of a program in the production stages? Both Markle (1964) and Lysaught and Williams (1963) indicate that errors in content accuracy, appropriateness, relevance, and writing style are not likely to be uncovered by students in trial situations, and suggest the use of external reviewers such as subject matter experts to assist with initial program editing. Again, Markle (1967) provides the most intricate and rigorous accounts of formative testing, suggesting that once content has been edited and reviewed to address the aforementioned factors, then one-on-one testing with learners in controlled settings should precede field trials involving larger numbers of learners. She insists that only frame-by-frame testing can provide accurate and reliable data not only about error rates, but also information pertaining to communication problems, motivational issues, and learning variables. Some design considerations may cross these three categories, such as the "size-of-step" issue (p. 121), which is both an instructional challenge as well as a motivational factor.

Once a program has been produced, many feel that it is the program producer's obligation to collect data regarding its effectiveness in the field (Glaser, Homme, & Evans, 1959; Lumsdaine, 1965; Markle, 1967). This contention was so compelling that a joint committee was formed from members representing the American Educational Research Association, the American Psychological Association, and the Department of Audiovisual Instruction (a division of the National Education Association). The report created by this Joint Committee on Programmed Instruction and Teaching Machines (1966) offers guidance to a variety of stakeholders regarding the evaluation of program effectiveness, including programmatic effectiveness data that prospective purchasers should seek, as well as guidelines for program producers and reviewers in their production of reports for the consumer. While the committee expresses the value inherent in one-on-one and small group testing, they place stronger emphasis on the provision of data from larger groups of students and repeated testing across groups to demonstrate the program's reliability and validity in effecting its intended outcomes.

In his description of considerations for program assessment, Lumsdaine (1965) is careful to point out the need to distinguish between the validation of a specific program and the validation of Programmed Instruction as an instructional method, a distinction that has continued through present-day evaluation concerns (Lockee, Moore, & Burton, 2001). Although evaluation and research may share common data collection approaches, the intentions of each are different, the former being the generation of product-specific information and the latter being concerned with the creation of generally applicable results, or principles for instruction (Lumsdaine, 1965).

20.4 RESEARCH ON PROGRAMMED INSTRUCTION

Skinner (1968b) lamented that many devices sometimes called *teaching machines* were designed and sold without true understanding of underlying pedagogical or theoretical aspects of their use. He noted that the design and functions of teaching machines and programmed instruction had not been adequately researched. Programmed Instruction was merely a way to apply technical knowledge of behavior to that of teaching. He called for additional experimental analysis that would look at behavior and its consequences, particularly in a *programmed* or sequenced instruction (Skinner, 1968b). The study of behavior through the analysis of reinforcement suggests a "new kind of educational research" (Skinner, 1968b, p. 414). Earlier research relied on measurement of mental abilities and comparisons of teaching methods and this led to a neglect of the processes of instruction. According to Skinner these types of comparisons and correlations are not as effective as results studied by manipulating variables and observing ensuring behavior. Moreover, in Skinner's view much of the earlier research was based upon "improvisations of skillful teachers" or theorists working "intuitively" and these types of studies had seldom "led direction to the design of improved practices (Skinner, 1968b, p. 415).

Skinner (1968b) stated that in dealing with research on Programmed Instruction, "No matter how important improvement in the students performance may be, it remains a by-product of specific changes in behavior resulting from the specific changes in the environment" (p. 415). With that said, there is a vast amount of literature on programmed instruction research that deals with student performance rather than specific changes in behavior and environment. Some proclaim a convincing array of evidence in its effectiveness; some results are provocative and unconvincing. Some research would qualify as good (in terms of methods, control, procedures) other research is no more than poor and contains repudiated techniques such as comparison studies. The 1950s and 1960s were the zenith of programmed instruction research in the literature. There are many compendiums and excellent reviews of this research. Some of these excellent sources of programmed instruction research follow. These included books by Stolurow (1961), Smith and Smith (1966), Lumsdaine and Glaser (1960), Glaser (1965), Taber et al. (1965), Ofiesh and Meirhenry (1964), Galanter (1959), and Hughes (1963) to name a few excellent research references. The research and evaluation issues and categorization of research components in program learning are many. This paper will look at general issues, research on teaching machines and devices, and variations and components of programs and programming. General issues include learning process and behavioral analysis, sole source of instruction, age level, subject matter properties and entering behavior and attitudes. Research summaries on teaching machines will review Pressey's self-instructional devices, military knowledge trainers, Skinner's teaching machines, and programmed books. Research on program variations will include programming variables response mode (such as linear and branching formats) prompts, step size, attitude, error rate, confirmation, and impact on age level.

20.4.1 A Disclaimer

The authors of this chapter, upon reviewing the research literature available found themselves in an ethical quandary. For the most part, the research conducted and published in this era of the zenith of programmed instruction use is generally poor. For example, many of the research studies conducted in the 1950s and 1960s were *comparison studies* that compared programmed materials and/or teaching machines with *conventional* or traditional methods. Despite their prevalence in this era's literature, most of these studies lack validity because the results cannot be generalized beyond the study that generated them, if at all. In addition, no program—machine-based or teacher-led, represents a whole category, nor do any two strategies differ in a single dimension. They cannot be compared because they differ in many ways (Holland, 1965). "The restrictions on interpretation of such a comparison arise from the lack of specificity of the instruction with which the instrument is compared" (Lumsdaine, 1962, p. 251). The ethical concern is that we have a large body of research that is for the most part ultimately not valid. It is also not reliable and could *not* meet the minimal standards of acceptable research. Unfortunately, much of this research was conducted by notable and experienced professionals in the field and published by the most reputable journals and organizations. Some of these problems were acknowledged early on by such researchers as Holland (1965, p. 107–109) and A. A. Lumsdaine (1962, p. 251). The authors of this chapter decided to proceed on the *buyer beware* theory. We present a limited sample of the literature addressing a variety of PI-related aspects, if for no other reason than to illustrate the breadth of the problems. For the most part, the research on programmed instruction began to die out in the early 1970s. This may have been due to editors finally realizing that the research products were poor or that the *fad* of programmed materials had slipped into history, likely the latter since it was replaced for the most part by equally flawed studies conducted on computer-assisted instruction.

Holland (1965), in recognizing the research concerns, felt that the "pseudo-experiments do not serve as a justified basis for decision" (p. 107). The answer is not to rely on this body (large) of research but to use evaluative measures, which tested against internal standards and requirements. As a result few generalizations will be made, but we will present the findings, summaries, and options of the original researchers. We will not critique the articles individually, but will allow the readers to judge for themselves.

20.4.2 Teaching Machines

20.4.2.1 Pressey's Machines. Pressey's self-instruction devices were developed to provide students with immediate feedback of results on knowledge after reading and listening to a lecture. Most of the research on Pressey's devices dealt with implementation and use of the results in order to develop a specific type of information to help the instructor change content and approach. Stolurow (1961) raised a question early on: when a programmed machine is used in conjunction with other means of instruction, which would be the cause of any effect? He felt it would be important to be able to judge how effective the programmed devices would be when used alone versus when they were used in conjunction with other types of instruction.

There was less concern about the problems of programming and sequencing in these machines (Stolurow, 1961). An example of research in this category was Peterson (1931) who evaluated Pressey's concepts with matched participants who were given objective pre- and posttests. The experimental group was given cards for self-checking their responses while the control group received no knowledge of feedback. In another version the participants were given a final test that was not the same as the posttests. In both situations the experimental group with knowledge of results scored higher than the control group. Little (1934) compared results from groups either using a testing machine, a drill machine, or neither (control group). Both experimental groups scored significantly higher than the control group. The group using the drill machine moved further ahead than did the test machine group. Other studies during the 1940s (as cited in Smith & Smith, 1966) used the concept of Pressey's devices. These concepts included punchboard quizzes, which gave immediate feedback and were found to significantly enhance learning with citizenship and chemistry content. Angell and Troyer, (1948), Jones and Sawyer (1949), Briggs (1947), and Jensen (1949) reported that good students using self-evaluation approaches with punch cards were able to accelerate their coursework and still make acceptable scores. Cassidy (1950), (a student of Pressey) in a series of studies on the effectiveness of the punchboard, reported that the immediate knowledge of results from this device provided significant increments in the learning of content. Pressey (1950) conducted a series of studies used punchboard concepts at The Ohio State University designed to test whether punchboard teaching machines could produce better learning performance by providing immediate knowledge of results and whether these beneficial effects are limited to a particular subject (Stolurow, 1961, p. 105). This series of studies lead to the following conclusions by Pressey and his associates as reported by Stolurow.

1. The use of the punchboard device was an easy way of facilitating learning by combining feedback, test taking, and scoring.
2. Test taking programs could be transformed to self-directed instruction programs.
3. When punchboards were used systematically to provide self-instruction, content learning was improved.
4. Automatic scoring and self-instruction could be achieved by the use of the punchboard.
5. The technique of providing learners with immediate knowledge of results via the punchboard could be used successfully in a variety of subjects. (1961).

Stephens (1960) found that using a Drum Tutor (a device used with informational material and multiple-choice questions and designed that students could not progress until the correct answer was made) helped a low-ability experimental group to score higher on tests than a higher ability group. This study confirmed Pressey's earlier findings that "errors were eliminated more rapidly with meaningful material and found that students learned more efficiently when they could correct errors immediately" (Smith & Smith, 1966, p. 249). These data also suggested that immediate knowledge of results made available early within the learning situation are more effective than after or later in the process (Stolurow, 1961). Severin (1960), another student of Pressey, used a punchboard testing procedure to compare the achievement of a learners forced to make overt responses versus those who were not required to make overt responses. No differences were reported. He concluded on short or easy tasks the automated overt devices were of little value. In an electrified version of the Pressey punchboard system, Freeman (1959) analyzed learner performance in a class of students who received reinforcement for a portion of the class and no reinforcement for another portion of time. He found no significant effects related to achievement; however, he indicated that in this study there were problems in the research design, including insufficient amount of reinforced opportunity, that test items were not identical to reinforced ones, and there was little attempt to program or structure the reinforced test materials (items). Freeman also noted that rapid gains in learning might not relate to better retention.

Holland (1959), in two studies on college students studying psychology using machine instruction, required one group of students to *space* their practice versus another group of students who had to mass their practice. He reported no significant differences as a result of practice techniques.

Stolurow (1961) suggested that studies on Pressey's machines, as a way of providing learners with immediate knowledge of results indicated that these machines could produce significant increments in learning, that learning by this method was not limited to particular subject areas and that the approach could be used with various types of learners. The effectiveness of having knowledge of results made available by these machines depended a great deal upon how systematic the material was programmed, the type of test to determine retention, and the amount of reinforced practice. Smith and Smith (1966) and Stolurow (1961) indicated that, based upon reviews of Pressey's earlier experiences, that there are positive outcomes of machine-based testing of programmed material. However, they also contended that the programmed machines may be more useful when used in connection with other teaching techniques. Pressey (1960), himself, states, "certainly the subject matter for automation must be selected and organized on sound basis. But the full potentialities of machines are now only beginning to be realized" (pp. 504–505). In reference to the effectiveness of programs on machine, Stolurow (1961) concluded that they are effective in teaching verbal and symbolic skills and for teaching manipulative skills.

Please note that there is a great overlap of the research on programmed machines and materials and of other approaches and variations. Additional programmed machine research is reviewed later in this section to illustrate points, concerns, and applications of other programming variables and research.

20.4.3 Military Knowledge Trainers

A major design and development effort in the use of automated self instruction machines was conducted by the U.S. Air Force, Office of Naval Research and by the Department of Defense during and after World War II. These development projects incorporated the concepts of Pressey's punchboard device in the forms of the Subject-Matter Trainer (SMT), the Multipurpose Instructional Problem Storage Device, the Tab-Item, and Optimal Sequence Trainer (OST), and the Trainer-Tester (see Briggs, 1960 for a description of these devices). These automated self instructional devises were designed to teach and test proficiency of military personnel. The Subject Matter Trainer (SMT) was modified to include several prompting, practice, and testing modes (Briggs, 1956, 1958). The emphasis of the SMT was to teach military personnel technical skills and content (Smith & Smith, 1966). Bryan and Schuster (1959) in an experiment found the use of the OST (which allowed immediate knowledge following a specific response) to be superior to regular instruction in a troubleshooting exam.

In an experimental evaluation of the Trainer-Tester and a military version of Pressey's punchboard, both devices were found to be superior to the use equipment mock-ups and of actual equipment for training Navy personnel in electronic troubleshooting (Cantor & Brown, 1956; Dowell, 1955). Briggs and Bernard (1956) reported that an experimental group using the SMT, study guides, and oral and written exams out performed the control group who used only the study guides and quizzes on a performance exam. However, the two groups were not significantly different on written tests. Both of these studies were related to the extent to which instruction provided by these machines was generalizeable or transferable. With respect to the effectiveness of these versions of teaching machines, these studies indicated that these programmed machines (SMT) can "be effective both for teaching verbal, symbolic skills which mediate performance and for teaching overt manipulative performance" (Stolurow, 1961, p. 115). Not all studies, however, reported superior results for the Subject Matter Trainer. He pointed out that these devices, which used military content and subjects generally, showed a consistent pattern of *rapid learning* at various ability levels and content and suggested that knowledge of results (if designed systematically) was likely to have valuable learning benefits.

20.4.4 Skinner's Teaching Machines

The research studies on Pressey's punchboard devices, and their military versions (e.g., SMT, OST, etc.), which incorporated many features of self-instruction and supported the concept that knowledge of results would likely have beneficial educational applications. However, the real impetus to self-instruction via machine and programmed instruction came from the theories and work of B.F. Skinner (e.g., 1954, 1958, 1961). Skinner's major focus was stating that self-instruction via programmed means should be in the context of reinforcement theory. He felt that Pressey's work was concerned "primarily with testing rather than learning and suggested that the important ideas about teaching machines and programmed instruction were derived from his analysis of operant conditioning" (Smith & Smith, 1966, p. 251). (See descriptions of these devices earlier in this chapter.) Skinner described his devices similar to Pressey's descriptions, including the importance of immediate knowledge of results. The major differences were that Pressey used a multiple-choice format and Skinner insisted upon *constructed* responses, because he felt they offered less chance for submitting wrong answers. Skinner's machines were designed to illicit overt responses. However, his design was modified several times over the years allowing more information to be presented and ultimately sacrificed somewhat the feature of immediate correction of errors. Skinner was most concerned about how the materials were programmed to include such concepts as overt response, size of steps, etc. As a result, much of the research was conducted on these programming components (concepts). These programming features included presenting a specific sequence of material in a linear, one-at-a-time fashion, requiring an overt response and providing immediate feedback to the response (Porter, 1958). Research on these components will be discussed later in this chapter. Much of the literature on Skinner's machines was in the form of descriptions of how these machines were used and how they worked (e.g., Holland, 1959; Meyer, 1959).

20.4.5 Crowder's Intrinsic Programming

Crowder (1959, 1960) (whose concepts were described earlier in this chapter) modified the Subject Matter Trainer to not only accommodate multiple choice questions, but to include his concept of *branching programming* in "which the sequence of items depends upon the response made by the student. Correct answers may lead to the dropping of certain items, or incorrect answers may bring on additional remedial material" (Smith & Smith, 1966, p. 273). Crowder's theories, like Skinner's were not machine specific. Much of the research was based around the various programmed aspects noted above. These programming aspects (variations) espoused by Crowder (1959, 1960) (e.g., large blocks of information, branching based upon response, etc.) will be also reviewed later in this chapter.

20.4.6 Programmed Instruction Variations and Components

As noted earlier, research on teaching machines and of programming components or program variations overlap to a great degree. Most teaching machines were designed to incorporate specific theories (e.g., Pressey—immediate knowledge of results in testing, and Skinner—overt responses with feedback in learning). Research on machines in reality became research on *program design and theory*. Because there was no general agreement on the best way to construct the machines or the programming approach much of the research deals with issues like type of programs, types of responses, size of steps, error rates, and the theoretical underpinnings of various approaches. The concept of programming refers to the way subject matter is presented, its sequence, its difficulty, and specific procedures designed into the program to enhance (theoretically) learning. It must be noted again that much of this research was conducted in the 1950s and 1960s and much of the research fell into the category of *comparison studies.* As such the reader should be weary of results and claims made by some of these researchers. The research summaries from this era and with its inherent problems provide no concrete answers or definitive results. They should, however, provide a feel for issues raised and potential insights about learning theory and their approaches.

20.4.7 Research on General Issues

20.4.7.1 Ability and Individual Differences. Glaser, Homme, and Evans (1959), suggested that individual differences of students could be important factor based upon previous research, which might affect program efficiency. Several questions arise under these assumptions: (1) Does student ability (or lack of) correlate with performance in a programmed environment, and (2) Does performance in a programmed environment *correlate* with *conventional instructional* methods and settings? Again, there appears to be no consensus in the results or the recommendations of the research.

Porter (1959) and Ferster and Sapon (1958) reported in separate studies that there was little or no correlation between ability level and achievement on programmed materials. Detambel and Stolurow (1956) found no relationship between language ability and quantitative subtests of ACE scores (American Council on Education Psychological Examination for College) and performance on a programmed task. Keisler (1959) matched two groups on intelligence, reading ability, and pretest scores, with the experimental group using a programmed lesson; the control group received no instruction. All but one of the experimental subjects scored higher after using the programmed materials.

Two groups of Air Force pilots were matched according to duties, type of aircraft, and "other" factors, with one group having voluntary access to a programmed self-tutoring game on a Navy Automatic Rater device. After two months the experimental group with voluntary access to the programmed materials showed significant improvement on items available with the game. The control group did not show significant improvement. However, there was no difference between the groups on items not included in the programmed materials. It was concluded that a self-instructional device would promote learning even in a voluntarily used game by matched subjects (Hatch, 1959).

Dallos (1976) in a study to determine the effects of anxiety and intelligence in learning from programmed learning found an interesting interaction on difficult programs. He reported that a high state of anxiety facilitated learning from the higher intelligence students and inhibited learning for low intelligence students.

Carr (1959) hypothesized that effective self instructional devices would negate differences in achievement of students of differing aptitudes. Studies by Porter (1959), and Irion and Briggs (1957), appeared to support this hypothesis as they reported in separate studies little correlation between intelligence and retention after using programmed devices. Carr (1959) suggested that the lack of relationship between achievement and intelligence and/or aptitude is because programmed instruction renders "learners more homogeneous with respect to achievement scores" (p. 561). Studies by Homme and Glaser (1959), and Evans, Glaser, and Homme (1959) tended to also support Carr's contention, while Keisler (1959) found students using machine instruction were more variable on achievement scores than the control group not using the programmed machines. Carr (1959) called for more study to determine the relationship between achievement and *normal predictors* with the use of programmed instruction.

20.4.8 User Attitude

Knowlton and Hawes (1962) noted, "that the pull of the future has always been slowed by the drag of the past" (p. 147). But, as there is a resistance to new technology, what proves valuable is thus too accepted. This statement appears to sum up the attitude toward programmed instruction in that perception of problems is due to lack of relevant information by the programmers and researchers.

Smith and Smith (1966) reported that the general reaction of learners towards programmed instruction at all levels including adult learners was very positive. This view was borne out by a number of studies gauging attitudes of learners toward programmed self-instruction. Stolurow (1963), in a study with retarded children using programmed machines to learn mathematics, found that these students, while apprehensive at first, later became engrossed and indicated they preferred using the machines rather than having traditional instruction. However, Porter (1959), in his earlier noted study, reported that there was no relationship among the gender of the student, the level of satisfaction with the programmed method, and achievement level. Students in a high school study revealed a view that was balanced between the use of programmed programs and conventional instruction (*First Reports on Roanoke Math Materials,* 1961). Eigen (1963) also reported a significant difference between attitudes use of programmed materials and other instruction of 72 male high school students in favor of the programmed instruction. Nelson (1967) found positive attitudes in student perceptions of programmed instruction in teaching music. Likewise, several studies on attitude were conducted in college classrooms. Engleman (1963) compared attitudes of 167 students using programmed and conventional instruction (lectures, labs, etc.) and reported that 28 percent indicated programmed materials were *absolutely essential,* 36 percent felt they were useful 90 percent of the time, 21 percent considered programmed materials useful 50 percent of the time, and 14 percent indicated that programmed materials were help only occasionally or not at all. Cadets at the Air Force Academy showed moderate enthusiasm as 80 percent indicated *enjoyment* in the programmed course, however, 60 percent preferred it to conventional teaching and suggested they learned with less effort (Smith, 1962). Several opinion studies were conducted in three colleges (Harvard, State College at Genesco, and Central Washington University) comparing attitudes of students using a programmed text, *The analysis of behavior,* (Holland & Skinner, 1961) and a textbook entitled *A textbook of psychology* (Hebb, 1958). The attitudes were overwhelming positive toward the programmed text (Naumann, 1962; VanAtta, 1961). Skinner and Holland (1960) reported that 78 percent of the students "felt they learned more form the machine than from the text" (p. 169). Banta (1963) reviewed similar attitude measures at Oberlin, University of Wisconsin, and Harvard and results were somewhat less favorable than the above study, but the Harvard students' attitude scores were similarly positive. Smith and Smith (1966) speculate that because the materials were developed at Harvard, there may have been a tendency to reflect their teachers' "enthusiasm and reacted in the expected manner" (p. 302). Roth (1963) also reported results of another college graduate students' opinion of the same Holland and Skinner text. All students liked it in the beginning, but only five did at the end of the study. Several objections noted that the program was "tedious," "repetitive," "mechanized," "non-thought provoking," and "anti-insightful" (Roth, 1963, p. 279–280). In a business setting at IBM, Hughes and McNamara (1961) reported that 87 percent of trainees liked programmed materials better than *traditional* instruction. Tobias (1969a, 1969b) provided evidence that teacher and user preferences for traditional devices are negatively related to achievement in programmed instruction. There have been a variety of studies dealing with student attitude toward various aspects of the programming variables. Jones and Sawyer (1949), in a study comparing attitudes of students using a programmed machine which provided self scoring and immediate knowledge of results versus a conventional paper answer sheet found 83 percent preferred the machine program over the paper answer sheet. Two studies (Eigen, 1963; Hough & Revsin, 1963) reported conflicting results on positive attitudes toward programmed machine and programmed texts. In a study concerning anxiety and intelligence when using difficult programmed instruction, Dallos (1974) found that participants with high anxiety, but lower intelligence had unfavorable view of the programmed instruction while the high intelligent, high anxiety participants had more favorable opinions of the program. Studies on attitude and learning effectiveness of programmed instruction have indicated that positive or negative attitudes toward programmed materials have little or no predictive value in determining learning effectiveness of these programs (Eigen, 1963; Hough & Revsin, 1963; Roe, Massey, Weltman, & Leeds, 1960; Smith & Smith, 1966). Smith and Smith (1966) indicated that these findings were not surprising because of other studies on general behavior have shown similar results (e.g., Brayfield & Crockett, 1955). "The apparent fact is that *general* attitude measures predict neither learning nor performance in a particular situation" (Smith & Smith, 1966, p. 304).

20.4.9 Programmed Instruction Compared to Conventional Instruction (Comparison Studies)

Much of the research on programmed machine and programmed instruction involved comparing programs to *conventional* or *traditional* instruction (whatever that was or is). This comparison technique was flawed from the beginning, but the results using this technique were used by many as proof the program was successful or was a failure, or was it *just as good as* the other form of instruction (incorrectly interpreting the *no significant difference* result).

Anytime one method of instruction is compared with another, several issues need to be kept in mind. Sometimes the comparisons are made between small groups with limited content and for relatively short time. Secondly, the novelty may effect operates in many cases generally supporting the new technique, e.g., programmed instruction. Thirdly, there are many, many uncontrolled factors operating all at once and any of these may affect the results of the study (Smith & Smith, 1966). This noted, in a review of 15 studies comparing programmed and conventional instruction, Silberman (1962) reported that nine favored programmed instruction and six indicated no significant difference in the two approaches. All 15 studies reported that the programmed approach took less time.

Several studies reported that when specific content was taught using programmed methods, time was saved with no decrease in achievement. All reported that instruction time was saved or the program-instruction completed requirements in less time than a conventional group (Hosmer & Nolan, 1962; Smith, 1962; Uttal, 1962; Wendt & Rust, 1962). In a study to compare a traditional instruction to a programmed method of teaching spelling in the third grade, the programmed group gained significantly better grade-equivalent scores than the control group by the end of the year (Edgerton & Twombly, 1962).

Hough (1962) compared machine programs to conventional instruction in a college psychology course where time was an additional factor. When quizzes were not announced the machine-instructed group scored significantly higher, but when quizzes were announced, there was no significant difference. Hough surmised that since the conventional group could study at home, whereas the machine group could not, the additional time available to the conventional group was a factor in these results.

Hartley (1966, 1972) reviewed 112 studies that compared programmed instruction (any variety) and *conventional instruction*. He concluded that there is evidence that programmed instruction is as good, or more effective than conventional instruction. In addition, Hamilton and Heinkel (1967) concurred in Harley's findings, which found in 11 of 12 studies that compared an instructor with a programmed lesson, an instructor alone, or a program alone, that an instructor with a program was the more effective choice. Hartley (1978) states "the results. . . . allow one to make the generalizations that many programs teach as successfully as many teachers and sometimes that they do this in less time" (p. 68). Falconer (1959) believed that it is an advantage for deaf children to use teaching machines where they traditionally require a large amount of individual instruction. He suggested that his data indicated that a teaching machine might be as effective as a teacher who had to spread his/her time over many students individually. Day (1959) compared a group using a *Crowder* style programmed book with that of conventional instruction. The experimental group that used the programmed book scored 20 percent higher and made one-fourth the wrong answers than the conventional instruction group over a half semester course. Goldstein and Gotkin (1962) reviewed eight experimental studies, which compared programmed text to programmed machines. Both versions were linear in nature. Goldstein and Gotkin reported no significant differences on several factors; posttest scores, time, and attitude across both presentation modes. (Four studies indicated the programmed texts used significantly less time than the machine version, however.) Other studies have shown no significant difference between automated instruction and traditionally taught classes or were equally effective modes of instruction (Goldberg, Dawson, & Barrett, 1964; Oakes, 1960; Tsai & Pohl, 1978). Similar no significant difference results were reported in studies with learning disabled students (e.g., Blackman & Capobianco, 1965; McDermott & Watkins, 1983; Price, 1963). Porter (1959) did report results showing that second and sixth graders progressed further in spelling achievement with programmed materials in less time than in a conventional classroom setting.

Silberman (1962) reviewed eight comparative studies to determine how best to present material in a self-instruction program, e.g., small step, prompting, overt response, branching, or repetition. He reported that there was no clear pattern of success and these cases showed that some treatments favored one method or another while other treatments favored the time-on-task factor. There were no significant differences across the programmed modes.

Eighth grade students of high ability were put into three groups, one used a linear program, one used a branching program, and the third was used as a control group (conventional instruction). Time available was constant across all groups. In a result unusual for this type of study, Dessart (1962) reported that the control group did significantly better than the experimental group using the branching approach. There was no significant difference between the conventional group and the linear group or between the linear and branching groups.

Stolurow (1963) studied the effect of programs teaching learning disabled children reading, vocabulary, and comprehension. Although, the results favored the programmed version over a traditional method, Stolurow recommended altering programs with conventional instruction. His recommendation was similar to others, which suggested a variety of methods may be more effective than only using one. Klaus (1961) reported on a comparison study dealing with 15 high school physics classes. Some classes had programmed materials available but not for mandatory use. The class having access to the programs had a substantial gain in criterion scores compared to the class without these materials available. After reviewing several studies, Alter and Silverman (1962) reported there were no significant differences in learning from the use of programmed materials or conventional texts. McNeil and Keisler (1962), Giese and Stockdale (1966), Alexander (1970), and Univin (1966) in studies comparing the two versions (programmed and conventional texts) also found the similar results of no significance across methods. However, in a number of studies using primarily retarded learners, the reported results of these comparison studies found the conventional instruction to be superior (Berthold & Sachs, 1974; McKeown, 1965; Richmond, 1983; Russo, Koegel, & Lovaas, 1978; Weinstock, Shelton, & Pulley, 1973). However, the programmed devices (particularly linear ones) have the advantage over teachers in a conventional setting who, in some cases, inadvertently skip over small ideas or points, which may need to be present for understanding. Some feel these programmed devices could solve this concern (Stolurow, 1961).

When program machines were studied as the sole source of instruction, Stolurow (1961) indicated in his review that both children and adults benefited from a programmed device. He stated, "these devices not only tend to produce performance which is freer of error than conventional methods of instruction, but also reduce the amount of instruction time required" (p. 135–136).

20.4.10 Programmed Variables (Essential Components)

During the early development of programmed instruction devices and materials many ideas were expressed on how best to present information, some based in theory (e.g., Skinner's work), others based on intuition, but little on actual research. Reviews of existing literature (e.g., Silberman, 1962) yielded no clear pattern of what programming criteria was effective in improving achievement. However, as time passed more studies and analyses of programming variables were conducted.

Program or programming variables are components that are essentially general in nature and can be associated with all types of programs. For an example, these variables can deal with theoretical issues such as the effect overt versus covert responses, the impact of prompting or no-prompting, size of steps, error rate, or the confirmation of results. Other issues indirectly related to the programming variables include user attitudes toward programs the mode of presentation (e.g., linear and branching) and program effectiveness. Illustrative results are provided from representative research studies.

20.4.11 Mode of Presentation

Various studies have been conducted comparing linear to branching programs, both in terms amount of learning and time saved in instruction. Coulson and Silberman (1960), and Roe (1962) found no significant differences in test scores between the two versions, but both found significant differences in terms of time taken to learn favoring branching programs. However, Roe (1962) did find that forward branching and linear programs were significantly faster (in terms of time saved) than backward branching. Mixed results were found in other studies, for example, Silberman, Melaragno, Coulson, and Estavan (1961) found no significant difference between the versions of presentation on achievement, but in the following study, Coulson, Estavan, Melaragno, and Silberman (1962) found that the branching mode was superior to a linear presentation.

Holland (1965), Leith, (1966), and Anderson (1967) reported no significant difference in learning between linear and branching programs when compared, and indicated this was generally the case with older or intelligent learners, "younger children using linear programs were more likely to receive higher test scores, although often these still took longer to complete than did branching ones" (Hartley, 1974, p. 284).

20.4.12 Overt Versus Covert Responses

One of Skinner's principles of programmed instruction is the necessity of overt responses. It appeared to be an important research concern to determine when it is advantageous to require overt or allow covert responses that could affect learning achievement. Are covert responses as effective as overt ones? This question has been a popular research topic. Overt responses require the student to *do something* (e.g., writing or speaking an answer, while covert requires *thinking* about or reading the material). Skinner's (1958) theory requires that a response should be overt (public) because if not overt, responses often ceased (Holland, 1965). Holland (1965) suggested that covert responses are not necessarily theoretical but also practical, because all aspects (in Skinner's view) of a program necessitate getting the correct answer. "Therefore, [a] measure of a program by not answering at all circumvents the characteristics which make it a program" (p. 93). Holland (1965) continued, indicating that several conditions must be met to determine the difference between overt and covert responses, namely, (1) program design must allow the student to answer correctly, and (2) the correct answer can only be attained after the appropriate steps in the program have been completed. Other researchers over the years have accepted this concept as important (e.g., Tiemann & Markle, 1990).

In reviews of research by Lumsdaine (1960, 1961), Feldhusen (1963), and Silberman (1962), all reported some mixed results, but the overall finding was that there was no difference in achievement between the overt or covert response groups. Results of several studies suggest that the use of overt responses was supported under some conditions (e.g., Briggs, Goldbeck, Campbell, & Nichols, 1962; Williams, 1963; Wittrock, 1963). Holland (1965) reported that when answers on a test are not contingent on important content, overt responding might not be effective. Otherwise, studies indicated a test advantage for students using overt responses. Goldbeck and Campbell (1962) found that the advantages of each type of response may vary with the difficulty of content. Additionally, several studies showed that overt responding in programmed instruction was beneficial over covert responses (Daniel & Murdock, 1968; Karis, Kent, & Gilbert, 1970; Krumboltz & Weisman, 1962; Tudor, 1995; Tudor & Bostow, 1991; Wittrock, 1963). Miller and Malott (1997) in a review of the literature on effectiveness of overt responses versus nonovert responses concluded that there was little benefit in requiring overt responses when additional learning-based incentives are present, but in situations where no incentives are present overt learning should improve learning.

A large number of other researchers found no significant difference between the effectiveness of programmed materials requiring overt responses and those using covert responses (Alter & Silberman, 1962; Csanyi, Glaser, & Reynolds, 1962; Daniel & Murdock, 1968; Goldbeck & Campbell, 1962; Goldbeck, Campbell, & Llewellyn, 1960; Hartman, Morrison, & Carlson, 1963; Kormandy & VanAtta, 1962; Lambert, Miller, & Wiley, 1962; Roe, 1960; Stolurow & Walker, 1962; Tobias, 1969a, 1969b, 1973; Tobais & Weiner, 1963). Shimamune (1992) and Vunovick (1995) found no significant difference between overt construction and discrimination responses and covert responses. However, in these studies extra credit (incentives) was given for test performance. Miller and Malott (1997) replicated Tudor's (1995) study and found that the no-incentives overt group produced greater improvement than did the covert responding group. This was also true for the incentive overt responding group as well. Their results did not support earlier studies (noted above) and concluded that overt responding was "robust enough phenomenon to occur even when an incentive is provided" (p. 500).

Evans et al. (1959) required two groups to use machine instruction except one group was required to answers overtly, the other group were required not to answer items overtly. They reported no significant difference in the approach, but the nonovert answering group took less time than the overt group. While the research reported primarily no significant difference between learners who wrote answers and thought about answers, Holland (1965), Leith (1966), and Anderson (1967) felt that there were situations in which overt answers were superior to covert answers. Hartley (1974) summarized these situations: (1) when young children were involved, (2) when materials were difficult or complex, (3) when programs were lengthy, and (4) when specific terminology was being taught. There is, however, evidence according to Glaser and Resnick (1972), and Prosser (1974) the mere questioning is important to learning, regardless of covert or overt response situations.

20.4.13 Prompting

Holland (1965) indicated that in a study of paired associates, prompting was defined as a response given prior to an opportunity to have an overt response, whereas when confirming the response item is given after the overt response. Several studies dealt with the advantages of prompting versus nonprompting in a program sequence. Cook and Spitzer (1960) and Cook (1961) reported a no significant difference between the two versions, and also indicated that overt responses were not necessary for better achievement. Angell and Lumsdaine (1961) concluded from the review several studies that programs should include both prompted and nonprompted components. Stolurow, Hasterok, and Ferrier (1960) and Stolurow, Peters, and Steinberg (1960) in preliminary results of a study reported the effectiveness of prompting and confirmation in teaching sight vocabulary to mentally retarded children. In an experiment comparing a partial degree of prompting (prompting on 3/4 of the trials) to a complete prompting (prompting on every trial) version, Angell and Lumsdaine (1961) found learning was significantly more efficient under the partial prompting condition and supported the results of Cook (1958) and Cook and Spitzer (1960).

20.4.14 Confirmation

There appears to be some controversy over the concept or interpretation of feedback, reinforcement, and confirmation. Skinner (1959) interpreted confirmation as a positive reinforcer in the operant conditioning model (Smith & Smith, 1966). Others have objected to this view suggesting that getting a student to perform a desired function for the first time is not addressed (Snygg, 1962). Lumsdaine (1962) suggested that program developers should be most interested in the manipulation of prompting cues, not manipulation of reward schedules. Smith and Smith (1966) indicated that in an operant conditioning situation the response and the reinforcement are constant while in programmed instruction the situations are continually changing.

Several studies compared programs with confirmation (after an overt answer, the correct answer is presented) to programs with no confirmation available. No significant difference was found in scores as a function of confirmation (Feldhusen & Birt, 1962; Holland, 1960; Hough & Revsin, 1963; Lewis & Whitwell, 1971; McDonald & Allen, 1962; Moore & Smith, 1961, 1962; Widlake, 1964). However, Meyer (1960), Angell (1949), and Kaess and Zeaman (1960) found significant advantages in answer confirmation. Suppes and Ginsberg (1962) found an overt correction after confirmation to be also effective. Krumboltz and Weisman (1962) in comparing continuous versus noncontinuous confirmation, reported neither had an effect on the test scores.

Repetition and review have been built into many programs. Some programs were designed to drop a question when it had been correctly answered. Because it was technically easier in 1960s to drop out a question after only one correct response rather than after additional responses, many programs were designed this way. However, Rothkopf (1960) did try to determine if there was any advantage to dropping questions out after two correct responses or any advantage to a version where none of the questions were dropped. He reported that the two methods were equally effective.

Scharf (1961) and Krumboltz and Weisman (1962) investigated several schedules of conformation and found no significant difference. However, Holland (1965) claimed even in the absence of significant results, that there was "enough suggestion of small differences so that the importance of confirmation cannot be discounted" (p. 91). Jensen (1949), Freeman (1959), and Briggs (1949) all reported that when there is a frequent, deliberate, and systematic effort to integrate the use of knowledge-of-results, learning shows a cumulative effect in a significant manner.

Hartley (1974) in his review and summary of programmed learning research on learner knowledge of results argued that immediate knowledge affected some learners more than others. In experiments "with low-ability learners and with programs with higher error rates, immediate knowledge of results was found to be helpful" (Holland, 1965; Anderson, 1967; Annett, 1969, as cited in Hartley, 1974, p. 284).

Although reinforcement, feedback, and confirmation are central issues to programmed instruction research, this area of research is incomplete and additional information concerning variables such as amount, schedule, and delay of reinforcement was missing. There appears to be no research that explains the problem of why confirmations are not always needed or why programs exhibiting the "pall effect" (boredom induced by the program) could promote learning (Rigney & Fry, 1961, p. 22).

20.4.15 Sequence

The basic structure of programmed machines and materials is a systematic progression of behavioral steps, which takes the student through complex subject matter with the intention of knowledge acquisition. One of Skinner's major tenants was the "construction of carefully arranged sequences of contingencies leading to the terminal performance which are the object of education" (Skinner, 1953, p. 169). This sequence of information and progressions in terms of "both stimulus materials displayed to the student and the way in which he interacts with and responds to them" are a fundamental issue of programmed learning research (Taber et al., 1965, p. 167).

Gavurin and Donahue (1960) compared a sequenced order of a program with a scrambled-order version as to the number of repetitions required for an errorless trial and on the number of errors to reach criterion. For both measures the sequenced order was significantly better. Hickey and Newton (1964) also found a significant difference in favor of original sequence to another unordered one. Hartley (1974) indicated that this suggested that the "analysis of structure must be very sophisticated indeed if it is to reveal useful differences in sequencing procedures" (p. 283). Roe, Case, and Roe (1962) found no significant difference post-test scores on a scrambled ordered versus a sequenced ordered program on statistics. However, using a longer form of the same program, Roe (1962) found significant advantages for the ordered sequences, on the number of student errors on the program and amount of time needed to complete the program.

Several research studies comparing *ordered* program sequences with nonlogical or random sequences have not supported Skinner's principle of ordered sequences (Duncan, 1971; Hamilton, 1964; Hartley & Woods, 1968; Miller, 1965; Neidermeyer, Browen, & Sulzen, 1968; Wager & Broaderick, 1974). However, Wodkte, Brown, Sands, and Fredericks (1968) found some evidence that the use of logical sequences for the lower ability learner was positive. Miller's (1969) study indicated that logical sequence appears to be the best in terms of overall effectiveness and efficiency. He felt it would be of value, however, to identify which levels of sequencing would be the most effective. In a review of several studies on logical sequencing, Hartley (1974) indicated that learners could tolerate "quite considerable distortions from the original sequence . . . and that the test results obtained are not markedly different from those obtained with the original program's so-called logical sequence" (p. 282). He stressed that these studies were conducted on short programs, however.

20.4.16 Size of Step

Size of step generally refers to the level of difficulty of the content or concepts provided in a frame. In addition, step size can mean, (1) amount of materials, for example, number of words in a frame, (2) difficulty as in error rate, and (3) number of items present (Holland, 1965). Thus, research in this category varies by "increasing or decreasing the number of frames to cover a given unit of instruction" (Smith & Smith, 1966, p. 311).

Using a programmed textbook with four levels of *steps* (from 30 to 68 items), four groups of students completed the same sequence of instruction, each group with a different number of steps. Evans et al. (1959) reported in that the group using smaller steps produced significantly fewer errors on both immediate and delayed tests. Likewise, Gropper (1966) found that larger the step size, the more errors were committed during practice. This finding was significant for lower ability students.

Smith and Moore (1962) reported in a study in which step size (step difficulty) and pictorial cues were varied in a spelling program, that no significant difference was found on achievement related to step size, but the larger step program took less time. Smith and Smith (1966) opined, "very small steps and over-cueing may produce disinterest"(p. 311). Balson (1971) also suggested that programmers could "increase the amount of behavioral change required of each frame" and thus increase the error rate, but not decrease achievement levels and also have a significant saving of time in learning (p. 205). Brewer and Tomlinson (1981) reported that except for brighter students, time spent on programmed instruction is not related to either improvement in immediate or delayed performance. Shay (1961) studied the relationship of intelligence (ability level) to step size. He reported relationship and indicted that the small steps were more effective (producing higher scores) at all ability levels.

Rigney and Fry (1961) summarized various studies and indicated that programs using very small (many components to a concept) could introduce a "pall effect" (Rigney & Fry, 1961, p. 22) in which boredom was inducted by the material, particularly with brighter students. These results were later supported by Briggs et al. (1962), Feldhusen, Ramharter, and Birt (1962), and Reed and Hayman (1962).

Coulson and Silberman (1959) compared three conditions on materials taught by machine: multiple-choice versus constructed responses, small steps versus large steps and branching versus no-branching presentation. This ambitious program's results indicated (1) that small steps (more items per concept) result in higher scores, but more training time, (2) the branching versions were not significantly different, but when time and amount of learning, the differences favored the branching version, and (3) there was no significant difference in the results of the type of response.

20.4.17 Error Rate

A major tenet in programmed instruction was presenting a sequence of instruction, which has a "high probability of eliciting desired performance" (Taber et al., p. 169). This sequence can sometimes be made too easy or too difficult. Error Rate is associated closely with size of step because of the codependence of the two. Skinner's (1954) thesis is that errors have no place in an effective program. They hinder learning. Others feel it is not necessarily an easy program (with few errors) that allows more learning but the program that involves and stimulates participation.

Again the results are mixed and generally dependent upon the situation. Studies by Keisler (1959), Meyer (1960), and Holland and Porter (1961) support the concept of low error rate. While Gagne′ and Dick (1962) found low correlations between error rate and learning others found the specific situation, topics, or content to be a major factor in this determination. Goldbeck and Campbell (1962) found overt responses were less effective in *easy* programs. Melaragno (1960) found that when errors occurred in close proximity in the program there was a negative outcome in achievement.

Several studies have looked at the question of the use of explanations for wrong answers. Bryan and Rigney (1956) and Bryan and Schuster (1959) found that explanations were particularly valuable with complex data. However, Coulson, Estavan, Melaragno, and Silberman (1962) found no difference in achievement between a group using linear programs with no knowledge of errors and a group using branching programs that provided explanations of errors. However, the students' level of understanding increased with explanation of errors.

20.4.18 Program Influence by Age or Level

Glaser, Reynolds, and Fullick (1963; as cited in Taber et al., 1965) conducted an extensive research study on program influence by grade level. This study was conducted within a school system using programmed materials at various grade levels, including first grade math, and fourth grade math subjects. The results were measured by program tests, teacher-made tests and by national standardized tests. One purpose of this study was to determine if very young students could work on and learn from programmed materials in a day-by-day plan. Glaser et al. reported that the students were successful in learning from the programmed materials, that students who completed the programs in the shortest time did not necessarily score the highest, that 95 percent of the students achieved 75 percent subject mastery, and 65 percent of the students at the fourth-grade level achieved 90 percent on the program and standardized test. While the researchers felt that the study was a success, they still felt that the role of the teacher *insured proficiency* by the students.

Many studies were conducted in the business and industry sector dealing with programmed instruction for training and reported significant instructional training success, a significant saving of time, or both (Hain & Holder, 1962; Hickey, 1962; Holt, 1963; Hughes & McNamara, 1961; Lysaught, 1962). A series of studies (e.g., Dodd, 1967; Evans, 1975; Mackie, 1975; Stewart & Chown, 1965) reviewed by Hartley and Davies (1978), concentrated on adults' use of programmed instruction. They concluded that there was no single *best* form (e.g., format, type) of programmed instruction, which is "appropriate for everyone at a given age doing a specific task" (p. 169). They also concluded that adults like and will work with programs longer than younger students and the more interaction built in, the more it is accepted by the adults.

20.4.19 Type of Response—Constructed vs. Multiple Choice

When errors (what some call negative knowledge) are made in a program in Skinner's (1958) view inappropriate behavior probably has occurred. Effective multiple-choice questions must contain opportunity for wrong answers and thus is out of place in the process of shaping behavior. Pressey (1960) and others claimed just the opposite, that "multiple-choice items are better *because* errors occur, permitting elimination of inappropriate behavior" (Holland, 1965, p. 86).

Several studies (Burton & Goldbeck, 1962; Coulson & Silberman, 1960; Hough, 1962; Price, 1962; Roe, 1960; Williams, 1963) compared constructed response and multiple-choice responses but found no significant differences. Fry (1960) however, found constructed responses to be the better approach.

Holland (1965) suggested a major advantage of programmed materials over other instructional methods is that they increase the probability of a correct answer. Nonprogrammed materials generally do not require an immediate answer or response, or the material is extraneous as far as the response is concerned. The more highly programmed materials have been demonstrated to be more effective in Holland's view.

20.4.20 Individual Versus Group Uses

Several studies have been conducted to assess the value of using programmed materials (various formats) in a group setting versus individual use. The results are mixed, Keisler and McNeil (1962) reported the findings of two studies using programmed materials, one showing a significant difference favoring the individual approach over the group approach. The second study found no significant difference in-group or individual approaches. Likewise, Feldhusen and Birt (1962) found no significance between individual and group approach. On the other hand, Crist (1967), reported positive results with group work with the use of programs over individual use.

20.4.21 Research Concerns

As noted earlier in the disclaimer, there has been much concern about the quality of research during the era of Programmed Instruction (Allen, 1971; Campaeu, 1974; Dick & Latta, 1970; Holland, 1965; Lockee et al., 2001; Lumsdaine, 1965; Moore, Wilson, & Armistead, 1986; Smith & Smith, 1966). There appears to be two major fundamental issues of concern, (1) poor research techniques and reporting, and (2) the preponderance of the comparison study. Smith and Smith (1966) noted several issues concerning PI research. These included:

1. Many of the comparisons used small groups, for limited subject areas and for very short study duration,
2. Because the concept of programmed instruction was relatively new in the 1950s and 1960s, the novelty effect tends to favor the new techniques, and
3. There are many uncontrolled effects apparent in many of the experiments, e.g., time.

Holland (1965) pointed out that no program or no conventional method is generic. Each program or teaching method is different in several ways (they have many, many characteristics that are uncounted for in many of these studies). The "adequacy of any *method* can be changed considerably by manipulating often subtle variables" (p. 107). Holland indicated that research on programmed learning was hampered by poor measures, test sensitivity, and experimental procedures. Campeau (1974), and Moldstad (1974) indicated rampant problems including lack of control, faulty reporting, small number of subjects, and a lack of randomization were present in many studies of this era (1950–1970). Stickell (1963) reviewed 250 comparative media studies conducted during the 1950s and 1960s and only 10 could be accurately analyzed. Most of the results were *uninterpretable*. His general assessment has great bearing on the era of programmed instruction research. The reliance on the comparison study for much of the research published during this time illustrates examples of faulty design and interpretation. Comparison studies assumed that each medium (e.g., programmed instruction) was unique and could or could not affect learning in the same way. This medium, in the researchers' views, was unique and had no other instructional attributes. These researchers give little thought to the medium's characteristics or those of the learners (Allen, 1971; Lockee et al., 2001). However, one must consider the question, "what are traditional instructional methods?" Most of these studies have used terms such as traditional or conventional instruction and have not specifically identified what these methods are. Research in which such variables are not properly identified should NOT be depended upon for valid results. Review of the many programmed instruction studies reveal incomplete, inaccurate, little or no descriptions of the treatments, methodology, and results (Moore, Wilson, & Armistead, 1986). Many of these studies, used very small samples (if they were actually samples), lacked randomization and misused and misinterpreted results. For example, a good number of this era's research studies used the statistical term, *no significant difference* to mean that variables were equally good or bad. Ask a poor question get a poor answer. Clearly any outcomes reported in these types of studies are invalid, but this fact did not stop many of the researchers, and for that matter, journal editors from misinterpreting or reporting these results (Levie & Dickie, 1973; Lockee et al., 2001).

20.4.22 Summary of Results

Stolurow (1961) felt that while research indicated that learners from learning disabled students to graduate students could effectively learn from programmed devices, additional research should continue and a systematic study of programming variables be developed.

Glaser (1960) noted early on in the era of programmed learning research that "present knowledge can scarcely fail be an improvement over anachronistic methods of teaching certain subjects by lecturing to large classes" (p. 30). Even at that time there was desire to deemphasize hardware and machines. But, that said, Glaser indicated that machines had the opportunity to offer tangibility over an existing instructional method alone and programmed machines had the opportunity to showcase the capabilities of reinforcement contingencies.

In early reviews of literature, Stolurow (1961) reported three general findings on programmed learning research: (1) a programmed machine can significantly enhance learning, (2) the advantages of programmed instruction are not limited by learning task or subject, and (3) teaching by programs are applicable to a variety of learners.

Stolurow (1961) in his summary of programmed learning literature stated that knowledge-of-results should be studied in more detail. He felt that knowledge-of-results would be more effective if given earlier in a learning situation and should be a bigger factor in programmed machine and material development.

While in Holland's (1965) view, the results of programmed variables have on paper supported the general theoretical foundations of programmed learning; the research has not "improved upon the principles because the studies have been limited to gross comparisons" (p. 92). He suggested future research, including the following aspects: (1) that the measuring and specifying of variables be more exact, and (2) that the research should be directed to improving existing procedures or developing new techniques. The *versus* statements found in many comparison study titles suggest *crude dichotomies,* without considering factors that might otherwise influence outcomes, such as other characteristics of the technology or the characteristics of the learner. "Consequently, a generalization of results is difficult since magnitudes of differences is important variables cannot be specified for either experimental materials or programs" (Holland, 1965, p. 92). That been said, Holland goes on to state that the research that to date (1966) supported the general principles of programming and in a paradoxal statement proclaimed that "it is perhaps comforting that comparison studies almost always show large advantages for programmed instruction" (p. 107). Holland (1965) stated that a contingent relationship between answer and content was important, that low error rate had received support, sequencing content was important and public, overt responses were important.

Hoko (1986) summarized his review of literature on the effects of automated instructional and traditional approaches, by indicating that each are unique and have specific potentials. He concluded, "the two should not be compared, but investigated, each for its own truths" (p. 18).

According to Smith and Smith (1966), the most valuable aspect of the program machine and instruction literature and research as that it provided "a new objective approach to the study of meaningful learning while at the same time provides new insights into how such learning occurs" (p. 326). While much of the research on programmed learning might be described as inconclusive, contradictory or even negative, there were important contributions. These contributions included focusing attention on reinforcement learning theory and possibly its shortcomings and thus opened the possibilities of new study and experimentation. Secondly, while not necessarily the norm, there were good researchers during this time that completed solid studies that did result in significant and meaningful results. This alone should indicate a need for more variability and research control to achieve real understandings of the programming theory and methods (Smith & Smith, 1966). Some authors and researchers felt that by the middle of the 1960s changes were in order and emphasis should (was) changing from emphasizing what the learner should do to what the programmer should do (Hartley, 1974). Some educators even felt that the psychology used to justify programmed instruction was becoming restrictive (Annett, 1969). Smith and Smith (1966) and Hartley and Davies (1978) tended to believe this earlier period of programming research started to shift from looking at program variables and learner needs to dealing with interactions with entire teaching and learning systems. Smith and Smith (1966) observed that this new emphasis on "systems study will not confine its efforts to evaluating specific machines or techniques, but will broaden its interests to include all types of classroom techniques and materials" (p. 326).

Computer-assisted instruction (CAI) and computer-based instruction (CBI) can be regarded as sophisticated extensions of programmed instruction theory and concept. Although some CBI research has been conducted within the context of programmed instruction, many of these studies have been conducted outside this context. Because of the many instructional possibilities that the computer can offer, many researchers consider it to be a separate field. This chapter's literature review dealt, for the most part, only with programmed instruction regarding theory and design. It should be noted that Programmed Instruction, CBI, and CAI have similar goals—to provide instruction, effectively, efficiently, and hopefully economically. It is evident that the foundations of computer-mediated instruction are based upon Programmed Instruction theory and research.

20.5 THE FUTURE OF PROGRAMMED INSTRUCTION

While trends in educational philosophy and learning theory have shifted away from behavioral sciences to more cognitive and constructivist approaches, these authors contend that Programmed Instruction has never really ceased to exist. Its influence is apparent in the instructional design processes that have continued to serve as the standards for our field (i.e., Dick, Carey, & Carey, 2000; Gagne, Briggs, & Wager, 1992; Gustafson & Branch, 1997, 2002; Kemp, Morrison, & Ross, 1998; Smith & Ragan, 1999). Recent literature regarding current trends in instructional design and technology indicates that while the systematic instructional design process has been embraced at varying levels across different venues (Reiser & Dempsey, 2002), its behavioral origins are still evident and notions of PI are found in existing practice. From the conduct of a needs assessment, to the establishment of clearly defined and measurable objectives, to the process of task analysis, the creation of assessment instruments and approaches that reflect the specified outcomes, the provision of opportunities for practice and feedback, to evaluation of the instructional program or product—all of these aspects of instructional design developed into the formation of a cohesive process as function of the Programmed Instruction movement. Perhaps the most prominent effect of the PI tradition on education as a whole is the convergence of the science of learning with the practice of teaching, the point originating from the first discussion of PI from Skinner (1954) himself in "The Science of Learning and the Art of Teaching".

As Januszewski (1999) indicates, "politics and political overtones are likely to be an undercurrent in any historical or conceptual study of educational technology" (p. 31). In the current era of political conservatism with a strong emphasis on accountability in education (no matter the organization or institution), the pendulum may likely swing back to favor this particular learning design. Though current trends in learning theory reflect less behavioral approaches to instruction (Driscoll, 2002), factors such as high-stakes testing in K–12 environments could promote a resurgence of aspects of PI, at least in terms of identification of measurable learning outcomes, mastery learning techniques, and the evaluation of instruction.

In "Programmed Instruction Revisited," Skinner (1986) proposed that the small computer is "the ideal hardware for Programmed Instruction" (p. 110). Extending the idea of the self-paced attribute of PI is the advent of the networked learning environment, making educational opportunities available anywhere and anytime. The revolution of the desktop computer, coupled with the diffusion of the Internet on a global scale, has provided access to unlimited learning resources and programs through distance education. In fact, perhaps the most prolific and long-standing example of computer-based PI, the PLATO (Programed Logic for Automatic Teaching Operation) program, has evolved into a Web-based learning environment that offers a variety of instructional programs to learners of all ages and walks of life, including incarcerated constituents. Created as a Programmed Instruction project at the University of Illinois in 1963. PLATO has continued development and dissemination since then, following the evolution of the computer and offering a range of computer-based curriculum that is unparalleled. While the primary design philosophy behind PLATO has shifted to feature more constructivist ideals (Foshay, 1998), it still maintains some of its PI foundations. For example, it preassesses learners to determine at what point they should engage in the program and if any remediation is necessary. Also, it tracks their progress, providing immediate feedback and guiding them to make accurate responses. Its use is also heavily evaluated. These features are its hallmarks, and the aspects of the program that have perpetuated throughout the aforementioned shifts in instructional philosophy and learning theory, giving credence to the influence of PI. While CBI, CAI, and now networked computer environments have expanded to support a greater variety of instructional approaches, Programmed Instruction still remains an effective and empirically validated possibility for the design of mediated instruction.

ACKNOWLEDGMENTS

The authors would like to thank Sara M. Bishop, Krista Terry, and Forrest McFeeters for their assistance in collecting the extensive array of historical materials necessary for the production of this chapter. We sincerely appreciate their help.

References

Alexander, J. E. (1970). *Vocabulary improvement methods, college level*. Knoxville, TN: Tennessee University. (ERIC Document Reproduction Service No. ED 039095)

Allen, W. H. (1971). Instructional media research, past, present and future. *AV Communication Review, 19*(1), 5–18.

Alter, M., & Silverman, R. (1962). The response in Programmed Instruction. *Journal of Programmed Instruction, 1,* 55–78.

Anderson, R. C. (1967). Educational psychology. *Annual Review of Psychology, 18,* 129–164.

Angell, D., & Lumsdaine, A. A. (1961). Prompted and unprompted trials versus prompted trials only in paired associate learning. In A. A. Lumsdaine (Ed.), *Student response in Programmed Instruction* (pp. 389–398). Washington, DC: National Research Council.

Angell, G. W. (1949). The effect of immediate knowledge of quiz results and final examination scores in freshman chemistry. *Journal of Educational Research, 42,* 391–394.

Angell, G. W., & Troyer, M. E. (1948). A new self-scoring test device for improving instruction. *School and Society, 67,* 84–85.

Annett, J. (1969). *Feedback and human behavior*. Harmondsworth, UK: Penguin.

Balson, M. (1971). The effect of sequence presentation and operant size on rate and amount of learning. *Programmed Learning and Educational Technology, 8*(3), 202–205.

Banta, T. J. (1963). Attitudes toward a programmed text: "The analysis of behavior" with a textbook of psychology. *Audiovisual Communication Review, 11,* 227–240.

Barlow, J. (1960). *Conversational chaining in teaching machine programs*. Richmond, IN: Earlham College.

Beck, J. (1959). On some methods of programming. In E. H. Galanter (Ed.), *Automatic teaching: The state of the art* (pp. 55–62). New York: John Wiley & Sons, Inc.

Benjamin, L. (1988). A history of teaching machines. *American Psychologist, 43*(9), 703–704.

Berthold, H. C., & Sachs, R. H. (1974). Education of the minimally brain damaged child by computer and by teacher. *Programmed Learning and Educational Technology, 11,* 121–124.

Blackman, L. S., & Capobianco, R. J. (1965). An evaluation of Programmed Instruction with the mentally retarded utilizing teaching machines. *American Journal of Mental Deficiency, 70,* 262–269.

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook 1: Cognitive domain.* New York: David McKay.

Brayfield, A. H., & Crockett, W. H. (1955). Employee attitudes and employee performance. *Psychology Bulletin, 52,* 396–424.

Brewer, I. M., & Tomlinson, J. D. (1981). SIMG: The effect of time and performance with modular instruction. *Programmed Learning and Educational Technology, 18*(2), 72–86.

Briggs, L. J. (1947). Intensive classes for superior students. *Journal of Educational Psychology, 38,* 207–215.

Briggs, L. J. (1956). *A troubleshooting trainer for the E-4 fire control system*. USAF Personnel Training Research Center. Report TN 56-94.

Briggs, L. J. (1958). Two self-instructional devices. *Psychological Reports, 4,* 671–676.

Briggs, L. J. (1960). Two self-instructional devices. In A. A. Lumsdaine & R. Glaser, (Eds.), *Teaching and programmed learning* (pp. 299–304). Washington, DC: National Education Association.

Briggs, L. J., & Bernard, G. C. (1956). Experimental procedures for increasing reinforced practice in training Air Force mechanics for an electronic system. In G. Finch & F. Cameron (Eds.), *Symposium on Air Force Human Engineering Personnel and Training Research* (pp. 48–58). Washington, DC: National Academy of Science, NRC.

Briggs, L. J., Goldbeck, R. A., Campbell, V. N., & Nichols, D. G. (1962). Experimental results regarding form of response, size of step, and individual differences in automated programs. In J. E. Coulson (Ed.), *Programmed learning and computer-based instruction* (pp. 86–98). New York: Wiley.

Brown, J. S., & VanLehn, K. (1980). Repair theory: A generative theory of bugs in procedural skills. *Cognitive Science, 4,* 389–426.

Burton, B. B., & Goldbeck, R. A. (1962). *The effect of response characteristics and multiple-choice alternatives on learning during programed instruction* (Technical report number 4). San Mateo, CA: American Institute for Research.

Bryan, G. L., & Rigney, J. W. (1956). *An evaluation of a method for shipboard training in operation knowledge USN of Naval Research* (Technical Report 18).

Bryan, G. L., & Schuster, D. H. (1959). *The effectiveness of guidance and explanation in troubleshooting training* (Technical Report No. 28). Electronics Personnel Research Group, University of Southern California.

Bullock, D. (1978). *Programmed Instruction.* Englewood Cliffs, NJ: Educational Technology Publications.

Cambre, M.A. (1981). Historical overview of formative evaluation of instructional media products. *Educational Communication and Technology Journal, 29,* 3–25.

Campeau, P. L. (1974). Selective review of the results of research on the use of audiovisual media to teach adults. *AV Communication Review, 22,* 5–40.

Cantor, J. H. & Brown, J. S. (1956). *An evaluation of the Trainer-Tester and Punch card-Tutor as electronics troubleshooting training aids* (Technical Report WAVTRADEVCEN 1257–2–1). United States Naval Training Device Center, Office of Naval Research, Port Washington, NY.

Carr, W. J. (1959). A functional analysis of self-instructional devices. In A. A. Lumsdaine & R. Glaser (Eds.), *Teaching machines and Programmed Instruction* (pp. 540–562). Washington, DC: National Education Association.

Casas, M. (1997). *The history surrounding the use of Skinnerian teaching machines and programmed instruction (1960–1970).* Unpublished master's thesis, Harvard University, Boston.

Cassidy, V. M. (1950). *The effectiveness of self-teaching devices in facilitating learning*. Unpublished dissertation. Columbus, OH: Ohio State University.

Cook, J. O. (1961). From audience participation to paired-associate learning and response analysis in paired-associate learning experiments. In A. A. Lumbsdaine (Ed.), *Student response in Programmed Instruction* (pp. 351–373). Washington, DC: National Research Council.

Cook, J. O., & Spitzer, M. E. (1960). Supplementing report: Prompting versus confirmation in paired-associate learning. *Journal of Experimental Psychology, 59,* 257–276.

Coulson, J. E., Estavan, D. P., Melaragno, R. J., & Silberman, H. F. (1962). Effects of branching in a computer controlled autoinstructional device. *Journal of Applied Psychology, 46,* 389–392.

Coulson, J. E., & Silberman, H. F. (1959). *Results of initial experiments in automated teaching* (Report number SP-73). Santa Monica, CA: Systems Development Corporation.

Coulson, J. E., & Silberman, H. F. (1960). Effects of three variables in a teaching machine. *Journal of Educational Psychology, 51,* 135–143.

Crist, R. L. (1967). Role of peer influence and aspects of group use of programmed materials. *AV Communication Review, 15,* 423–434.

Cronbach, L. J. (1963). Course improvement through evaluation. *Teachers' College Record, 64,* 672–683.

Crowder, N. A. (1959). Automatic tutoring by means of intrinsic programming. In E. Galanter (Ed.), *Automatic teaching: The state of the art* (pp. 109–116). New York: Wiley.

Crowder, N. A. (1960). Automatic tutoring by intrinsic programming. In A. A. Lumsdaine & R. Glaser, (Eds.), *Teaching machines and programmed learning* (pp. 286–298). Washington, DC: National Education Association.

Crowder, N. A. (1964). On the differences between linear and intrinsic programming. In J. P. DeCecco (Ed.), *Educational technology* (pp. 142–151). New York: Rinehart & Winston.

Csanyi, A. P., Glaser, R., & Reynolds, J. H. (1962). Programming method and response mode in a visual-oral response task. In *Investigations of learning variables in Programmed Instruction.* Pittsburgh, PA: University of Pittsburgh.

Dale, E. (1967) Historical setting of Programmed Instruction. In P. C. Lange (Ed.), *Programed Instruction: The sixty-sixth yearbook of the National Society for the Study of Education.* Chicago, National Society for the Study of Education. 28–54.

Dallos, R. (1976). The effects of anxiety and intelligence on learning from Programmed Instruction*. Programmed Learning and Educational Technology, 13*(2), 69–76.

Daniel, W. J,. & Murdock, P. (1968). Effectiveness of learning from a programmed text compared with conventional text covering the same material. *Journal of Educational Psychology, 59,* 425–431.

Day, J. H. (1959). Teaching machines. *Journal of Chemical Education, 36,* 591–595.

Dessart, D. J. (1962). A study in programmed learning. *School Science & Mathematics, 62,* 513–520.

Detambel, M. H., & Stolurow, L. M. (1956). Stimulus sequence and concept learning. *Journal of Experimental Psychology, 51,* 34–40.

Dick, W., Carey, L., & Carey, J. O. (2000). *The systematic design of instruction* (5th ed.). Reading, MA: Addison Wesley.

Dick, W., & Latta, R. (1970). Comparative effects of ability and presentation mode in computer-assisted instruction and Programmed Instruction. *AV Communication Review, 18*(1), 33–45.

Dodd, B. T. (1967). A study in adult retraining: The gas man. *Occupational Psychology, 41,* 143.

Dowell, E. C. (1955). *An evaluation of Trainer-Testers, TA and D, Air Force Technical Training* (Report No. 54–28). Keesler Air Force Base, MS.

Driscoll, M. P. (2002). Psychological foundations of instructional design. In R. A. Reiser & J. A. Dempsey (Eds.), *Trends and issues in instructional design and technology* (pp. 57–69). Upper Saddle River, NJ: Merrill/Prentice Hall.

Duncan, K. D. (1971). Fading of prompts in learning sequences. *Programmed Learning and Educational Technology, 8*(2), 111–115.

Edgerton, A. K., & Twombly, R. M. (1962). A programmed course in spelling. *Elementary School Journal. 62,* 380–386.

Eigen, L. D. (1963). High school student reactions to Programmed Instruction. *Phi Delta Kappan, 44,* 282–285.

Engelmann, M. D. (1963). Construction and evaluation of programmed materials in biology classroom use*. American Biology Teacher, 25,* 212–214.

Evans, L. F. (1975). Unconventional aspects of educational technology in an adult education program. In E. F. Evans & J. Leedham (Eds.), *Aspects of educational technology IX,* London: Kogan Page.

Evans, J. L., Glaser, R., & Homme, L. E. (1959). *A preliminary investigation of variation in properties of verbal learning sequences of the teaching machine type*. Paper presented at the meeting of the Eastern Psychological Association, Atlantic City, NJ.

Evans, J. L., Glaser, R., & Homme, L. E. (1960). A preliminary investigation of variation in properties of verbal learning sequences of the teaching machine type. In A. A. Lumsdaine & R. Glaser (Eds.), *Teaching machines and programmed learning: A sourcebook* (446–451). Washington, DC: National Education Association.

Evans, J. L., Homme, L. E., & Glaser, R. (1962). The ruleg (rule-example) system for the construction of programmed verbal learning sequences. *Journal of Educational Research, 55,* 513–518.

Falconer, G. A. (1959). *A mechanical device for teaching word recognition to young deaf children.* Unpublished doctoral dissertation, University of Illinois, Champaign, IL.

Feldhusen, J. F. (1963). Taps for teaching machines. *Phi Delta Kappan, 44,* 265–267.

Feldhusen, J. F., & Brit, A. (1962). A study of nine methods of presentation of programmed learning materials. *Journal of Educational Research, 55,* 461–465.

Feldhusen, J. F., Ramharter, H., & Birt, A. T. (1962). The teacher versus programmed learning. *Wisconsin Journal of Education, 95*(3), 8–10.

Ferster, C. D., & Sapon, S. M. (1958). An application of recent developments in psychology to the teaching of German. *Harvard Educational Review, 28,* 58–69.

First Reports on Roanoke Math Materials. (1961). *Audiovisual Instruction, 6,* 150–151.

Foshay, R. (1998). *Instructional philosophy and strategic direction of the PLATO system*. Edina, MN: PLATO, Inc. (ERIC Document Reproduction Service ED 464 603.)

Freeman, J. T. (1959). The effects of reinforced practice on conventional multiple-choice tests. *Automated Teaching Bulletin, 1,* 19–20.

Fry, E. B. (1960). A study of teaching-machine response modes. In A. A. Lumsdaine & R. Glaser (Eds.), *Teaching machines and programmed learning* (pp. 469–474). Washington, DC: National Education Association.

Fry, E. B. (1961). Programming trends. *Audiovisual Instruction, 6,* 142–143.

Gagne, R.M. (1956). *The conditions of learning*. New York: Holt, Rinehart, and Winston.

Gagne′, R. M., Briggs, L. J., & Wager, W. W. (1992). *Principles of instructional design* (4th ed.). New York: Harper Collins.

Gagne′, R. M., & Dick, W. (1962). Learning measures in a self-instructional program in solving equations. *Psychology Reports, 10,* 131–146.

Galanter, E. H. (1959). *Automatic teaching: The state of the art*. New York: Wiley & Sons.

Gavurin, E. I., & Donahue, V. M. (1960). *Logical sequence and random sequence teaching-machine programs*. Burlington, MA: RCA.

Giese, D. L., & Stockdale, W. A. (1966). Comparing an experimental and conventional method of teaching linguistic skills. *The General College Studies, 2*(3), 1–10.

Gilbert, T. F. (1962). Mathetics: The technology of education. *Journal of Mathetics, I,* 7–73.

Glaser. R. (1960). Christmas past, present, and future: A review and preview. In A. A. Lumsdaine & R. Glaser (Eds.), *Teaching machines and programmed learning: A sourcebook* (pp. 23–31). Washington, DC: National Education Association.

Glaser, R. (Ed.), (1965). *Teaching machines and programmed learning II*. Washington, DC. National Education Association.

Glaser, R., Homme, L. E., & Evans, J. F. (1959, February). *An evaluation of textbooks in terms of learning principles*. Paper presented at the meeting of the American Educational Research Association, Atlantic City, NJ.

Glaser, R., & Klaus, D.J. (1962). Proficiency measurement: Assessing human performance. In R. M. Gagne (Ed.), *Psychological principles in system development*. New York: Holt, Rinehart, and Winston.

Glaser, R., & Resnick, L. B. (1972). Instructional psychology. *Annual Review of Psychology, 23,* 207–276.

Glaser, R., Reynolds, J. H., & Fullick, M. G. (1963). *Programmed Instruction in the intact classroom*. Pittsburgh, PA: University of Pittsburgh.

Goldbeck, R. A., & Campbell, V. N. (1962). The effects of response mode and response difficulty on programmed learning. *Journal of Educational Psychology, 53,* 110–118.

Goldbeck, R. A., Campbell, V. N., & Llewellyn, J. E. (1960). *Further experimental evidence in response modes in automated instruction.* Santa Barbara, CA: American Institute for Research.

Goldberg, M. H., Dawson, R. I., & Barrett, R. S. (1964). Comparison of programmed and conventional instructional methods. *Journal of Applied Psychology, 48,* 110–114.

Goldstein, L. S., & Gotkin, L. G. (1962). A review of research: Teaching machines vs. programmed textbooks as presentations modes. *Journal of Programmed Instruction, 1,* 29–36.

Green, E. (1967). The process of instructional programming. In P. C. Lange (Ed.), *Programed Instruction: The sixty-sixth yearbook of the National Society for the Study of Education* (pp. 61–80). Chicago, National Society for the Study of Education.

Gropper, G. L. (1966). *Programming visual presentations for procedural learning. Studies in televised instruction*. Pittsburgh, PA: American Institute for Research in Behavioral Sciences.

Gustafson, K., & Branch, R. M. (1997). *Survey of instructional development models* (3rd Edition). Syracuse, NY. (ERIC Clearinghouse on Information & Technology.)

Gustafson, K. & Branch, R. M. (2002). What is instructional design? In R. A. Reiser & J. A. Dempsey (Eds.), *Trends and issues in instructional design and technology* (pp. 16–25). Upper Saddle River, NJ: Merrill/Prentice Hall.

Hain, K. H., & Holder, E. J. (1962). A case study in Programmed Instruction. In S. Margulies & L. D. Eigen (Eds.), *Applied Programmed Instruction* (pp. 294–297). New York: John Wiley & Sons.

Hamilton, N. R. (1964). Effect of logical versus random sequencing of items in auto-instructional program under two conditions of covert response. *Journal of Educational Psychology, 55,* 258–266.

Hamilton, R. S., & Heinkel, O. A. (1967). *An evaluation of Programmed Instruction*. San Diego, CA: San Diego City College. (ERIC Document Reproduction Service No. ED 013619.)

Hartley, J. (1966). Research report. *New Education, 2*(1), 29.

Hartley, J. (Ed.) (1972). *Strategies for Programmed Instruction*. London: Butterworths.

Hartley, J. (1974). Programmed Instruction 1954–1974: A review. *Programmed Learning and Educational Technology, 11,* 278–291.

Hartley, J. (1978). *Designing instructional text*. New York: Nichols Publishing.

Hartley, J., & Davies, I. (1978). Programmed learning and educational technology. In M. Howe (Ed.), *Adult learning: Psychological research and applications* (pp. 161–183). New York: Wiley and Sons.

Hartley, J. E., & Woods, P. M. (1968). Learning poetry backwards. *National Society for Programmed Instruction Journal, 7,* 9–15.

Hartman, T. F., Morrison, B. A., & Carlson, M. E. (1963). Active responding in programmed learning materials. *Journal of Applied Psychology, 47,* 343–347.

Hatch, R. S. (1959). *An evaluation of a self-tutoring approach applied to pilot training* (Technical Report 59–310, p. 19). USAF Wright Air Development Center.

Hebb, D. O. (1958). *A Textbook of Psychology*. New York: Saunders.

Heinich, R. (1970). *Technology and the management of instruction* (Association for Educational Communications and Technology Monograph No. 4). Washington, DC: Association for Educational Communications and Technology.

Hickey, A. E. (1962). Programmed Instruction in business and industry. In S. Margulies & L. D. Eigen (Eds.), *Applied Programmed Instruction* (pp. 282–293). New York: John Wiley & Sons.

Hickey, A. E. & Newton, J. M. (1964). *The logical basis of teaching: The effect of subconcept sequence on learning.* Newbury Port, MA: Entelek Inc.

Hoko, J. A. (1986, February). What is the scientific value of comparing automated and human instruction? *Educational Technology, 26*(2) 16–19.

Holland, J. G. (1959). A teaching machine program in psychology. In E. Galanter (Ed.), *Automatic teaching: The state of the art* (pp. 69–82). New York: John Wiley & Sons.

Holland, J. G. (1960). Design and use of a teaching machine and program. *Teacher College Record, 63,* 56–65.

Holland, J. G. (1965). Research on programming variables. In R. Glaser (Ed.), *Teaching machines and programed learning, II* (pp. 66–177). Washington, DC: National Education Association.

Holland, J. G., & Porter, D. (1961). The influence of repetition of incorrectly answered items in a teaching-machine program. *Journal of Experimental Analysis Behavior, 4,* 305–307.

Holland, J. G., & Skinner, B. F. (1961). *The analysis of behavior*. New York: McGraw-Hill.

Holt, H. O. (1963). An exploratory study of the use of a self-instructional program in basic electricity. In J. L. Hughes (Ed.), *Programmed learning: A critical evaluation* (pp. 15–39). Chicago, IL: Educational Methods.

Homme, L. E., & Glaser, R. (1959, February). *Problems in programming verbal learning sequences*. Paper presented at the meeting of the American Psychological Association, Cincinnati, OH.

Hosmer, C. L., & Nolan, J. A. (1962). Time saved by a tryout of automatic tutoring. In S. Margulies & L. D. Eigen (Eds.), *Applied Programmed Instruction* (pp. 70–72). New York: John Wiley & Sons.

Hough, J. B. (1962). An analysis of the efficiency and effectiveness of selected aspects of machine instruction. *Journal of Educational Research, 55,* 467–471.

Hough, J. B., & Revsin, B. (1963). Programmed Instruction at the college level: A study of several factors influencing learning. *Phi Delta Kappan, 44,* 286–291.

Hughes, J. E. (Ed.), (1963). *Programmed learning: A critical evaluation*. Chicago, IL: Educational Methods.

Hughes, J. L., & McNamara, W. J. (1961). A comparative study of programmed and conventional instruction in industry. *Journal of Applied Psychology, 45,* 225–231.

Irion, A. L., & Briggs, L. J. (1957). *Learning task and mode of operation and three types of learning tasks on the improved Subject Matter Trainer*. Lowry Air Force Base, Air Force Personnel and Training Research Center, CO.

Januszewski, A. (1999). Forerunners to educational technology. In R. M. Branch & M. A. Fitzgerald (Eds.), *Educational media and technology yearbook, Volume 24* (pp. 31–42). Englewood, CO: Libraries Unlimited.

Jensen, B. T. (1949). An independent-study laboratory using self-scoring tests. *Journal of Educational Research, 43,* 134–137.

Joint Committee on Programmed Instruction and Teaching Machines. (1966). Recommendations on reporting the effectiveness of Programmed Instruction materials. *AV Communication Review, 14*(1), 117–123.

Jones, H. L., & Sawyer, M. O. (1949). A new evaluation instrument. *Journal of Educational Research, 42,* 381–385.

Kaess, W., & Zeaman, D. (1960). Positive and negative knowledge of results on a Pressey-type punchboard. *Journal of Experimental Psychology, 60,* 12–17.

Karis, C., Kent, A., & Gilbert, J. E. (1970). *The interactive effect of responses per frame response mode and response confirmation on intra-frame S-R association strength. Final Report.* Boston, MA: Northwestern University. (ERIC Document Reproduction Service No. ED 040591.)

Keisler, E. R. (1959). The development of understanding in arithmetic by a teaching machine. *Journal of Educational Psychology, 50,* 247–253.

Keisler, E. R., & McNeil, J. D. (1962). Teaching science and mathematics by autoinstruction in the primary grades: An experimental strategy in curriculum development. In J. E. Coulson (Ed.), *Programmed learning and computer instruction* (pp. 99–112). New York: Wiley.

Kemp, J., Morrison, G., & Ross, S. (1998). *Designing effective instruction* (2nd ed.). New York: Merrill.

Klaus, D. J. (1961). Programming: A re-emphasis on the tutorial approach. *Audiovisual Instruction, 6,* 130–132.

Knowlton, J., & Hawes, E. (1962). Attitude: Helpful predictor of audiovisual usage. *AV Communication Review, 10*(3), 147–157.

Kormondy, E. J., & VanAtta, E. L. (1962). Experiment in self-instruction in general biology. *Ohio Journal of Science, 4,* 4–10.

Krumboltz, J. D., & Weisman, R. G. (1962). The effect of covert responding to Programmed Instruction on immediate and delayed retention. *Journal of Educational Psychology, 53,* 89–92.

Lambert, P., Miller, D. M., & Wiley, D. E. (1962). Experimental folklore and experimentation: The study of programmed learning in the Wauwatosa Public Schools. *Journal of Educational Research, 55,* 485–494.

Lange, P. C. (1967). Future developments. In P. C. Lange (Ed.), *Programed instruction: The sixty-sixth yearbook of the National Society for the Study of Education* (pp. 284–326). Chicago, National Society for the Study of Education.

Leith, G. O. M. (1966). *A handbook of programmed learning* (2nd ed.). Educational Review Occasional Publication Number 1. Birmingham, UK: University of Birmingham.

Lewis, D. G., & Whitwell, M. N. (1971). The effects of reinforcement and response upon programmed learning in mathematics. *Programmed Learning and Educational Technology, 8*(3), 186–195.

Little, J. K. (1934). Results of the use of machines for testing and for drill upon learning educational psychology. *Journal of Experimental Education, 3,* 45–49.

Lockee, B. B., Moore, D. M., & Burton, J. K. (2001). Old concerns with new distance education research. *Educause Quarterly, 24,* 60–62.

Lumsdaine, A. A. (1960). Some issues concerning devices and programs for automated learning. In A. A. Lumsdaine & R. Glaser (Eds.), *Teaching machines and programmed learning* (pp. 517–539). Washington, DC: National Education Association.

Lumsdaine, A. A. (1961). *Student response in Programmed Instruction.* Washington, DC: National Research Council.

Lumsdaine, A. A. (1962). Experimental research on instructional devices and materials. In R. Glaser (Ed.), *Training research and education* (pp. 247–294). Pittsburgh, PA: University of Pittsburgh.

Lumsdaine, A. A. (1965). Assessing the effectiveness of instructional programs. In R. Glaser (Ed.), *Teaching machines and programmed learning, II: Data and directions* (pp. 267–320). Washington, DC: National Education Association.

Lumsdaine, A. A., & Glaser, R. (Eds.) (1960). *Teaching machines and programmed learning*. Washington, DC: National Education Association.

Lysaught, J. P. (1962). Programmed learning and teaching machines in industrial training. In S. Margulies & L.D. Eigen, (Eds.), *Applied Programmed Instruction* (pp. 23–43). New York: Wiley.

Lysaught, J. P., & Williams, C. M. (1963). *A guide to programmed instruction*. New York: John Wiley and Sons.

Mackie, A. (1975). Consumer-oriented programmed learning in adult education. In L. F. Evans & J. Leedham (Eds.), *Aspects of educational technology, IX,* London: Kogan Page.

Mager, R. F. (1962). *Preparing objectives for programmed instruction*. Belmont, CA: Fearon.

Malpass, L. F., Hardy, M. M., & Gilmore, A. S. (1964). Automated instruction for retarded children. *American Journal of Mental Deficiency, 69,* 405–412.

Markle, S. M. (1964). *Good frames and bad: A grammar of frame writing.* New York: John Wiley & Sons.

Markle, S. M. (1967). Empirical testing of programs. In P. C. Lange (Ed.), *Programed Instruction: The sixty-sixth yearbook of the National Society for the Study of Education* (104–138). Chicago: National Society for the Study of Education.

McDermott, P. A., & Watkins, W. W. (1983). Computerized vs. conventional remedial instruction for learning disabled pupils. *Journal of Special Education, 17,* 81–88.

McDonald, F. J., & Allen, D. (1962). An investigation of presentation response and correction factors in Programmed Instruction. *Journal of Educational Research, 55,* 502–507.

McKeown, E. N. (1965). A comparison of the teaching of arithmetic in grade four by teaching machine, programmed booklet, and traditional methods. *Ontario Journal of Educational Research, 7,* 289–295.

McNeil, J. D., & Keisler, E. R. (1962). Questions versus statements as stimuli to children's learning. *Audiovisual Communication Review, 10,* 85–88.

Mechner, F. (1967). Behavioral analysis and instructional sequencing. In P. C. Lange (Ed.), *Programed Instruction: The sixty-sixth yearbook of the National Society for the Study of Education* (pp. 81–103). Chicago: National Society for the Study of Education.

Melaragno, R. J. (1960). Effect of negative reinforcement in an automated teaching setting. *Psychological Reports, 7,* 381–384.

Merrill, M. D. (1971). Components of a cybernetic instructional system. In M. D. Merrill (Ed.), Instructional design: Readings (pp. 48–54). Englewood Cliffs, NJ: Prentice-Hall, Inc.

Meyer, S. R. (1959). A program in elementary arithmetic: present and future. In E. Galanter (Ed.), *Automatic teaching: The state of the art* (pp. 83–84). New York: John Wiley & Sons.

Meyer, S. R. (1960). Report on the initial test of a junior high-school vocabulary program. In A.A. Lumsdaine & R. Glaser (Eds.), *Teaching machines and programmed learning* (pp. 229–246). Washington, DC: National Education Association.

Miller, H. R. (1965, April). *An investigation into sequencing and prior information variables in a programmed evaluation unit for junior high school mathematics.* Paper presented at the meeting of the Department of Audiovisual Instruction. Milwaukee, WI.

Miller, H. R. (1969). Sequencing and prior information in linear Programmed Instruction. *AV Communication Review, 17*(1), 63–76.

Miller, M. L., & Malott, R. W. (1997). The importance of overt responding in Programmed Instruction even with added incentives for learning. *Journal of Behavioral Education, 7*(4), 497–503.

Miller, R. B. (1953). *A method for man–machine task analysis* (Tech. Rep. No. 53–137). Wright–Patterson Air Force Base, Ohio: Wright Air Development Center.

Miller, R. B. (1962). Analysis and specification of behavior for training. In R. Glaser (Ed.), *Training research and education*. Pittsburgh: University of Pittsburgh.

Moldstad, J. A. (1974). Selective review of research studies showing media effectiveness: A primer for media directors. *AV Communication Review, 22*(4), 387–407.

Moore, D. M., Wilson, L., & Armistead, P. (1986). Media research: A graduate student's primer. *British Journal of Educational Technology, 3*(17), 185–193.

Moore, J. W., & Smith, W. I. (1961). Knowledge of results of self-teaching spelling. *Psychological Reports, 9,* 717–726.

Moore, J. W., & Smith, W. I. (1962). A comparison of several types of immediate reinforcement. In W. Smith & J. Moore (Eds.), *Programmed learning* (pp. 192–201). New York: VanNostrand.

Naumann, T. F. (1962). A laboratory experience in programmed learning for students in educational psychology. *Journal Programmed Instruction, 1,* 9–18.

Neidermeyer, F., Brown, J., & Sulzen, R. (1968, March). *The effects of logical, scrambled and reverse order sequences on the learning of a series of mathematical tasks at the math grade level.* Paper presented at the meeting of the California Educational Research Association, Oakland, CA.

Nelson, C. B. (1967). *The effectiveness of the use of programmed analysis of musical works on students' perception of form. Final report.* SU New York at Cortland, NY.

Oakes, W. F. (1960). The use of teaching machines as a study aid in an introductory psychology course. *Psychological Reports, 7,* 297–303.

Ofiesh, G. D., & Meirhenry (Eds.). (1964). *Trends in Programmed Instruction,* Washington, DC: National Education Association.

Orey, M. A., & Burton, J. K. (1992). The trouble with error patterns. *Journal of Research on Computers in Education, 25*(1), 1–15.

Peterson, J. C. (1931). The value of guidance in reading for information. *Trans. Kansas Academy of Science, 34,* 291–296.

Porter, D. (1958). Teaching machines. *Harvard Graduate School Education Association Bulletin, 3,* 1–5.

Porter, D. (1959). Some effects of year long teaching machine instruction. In E. Galanter (Ed.), *Automatic teaching: The state of the art* (pp. 85–90). New York: John Wiley & Sons.

Posser, G. V. P. (1974). The role of active questions in learning and retention of prose material. *Instructional Science, 2,* 241–246.

Pressey, S. L. (1926). A simple apparatus which gives tests and scores—and teaches. *School and Society, 23,* 373–376.

Pressey, S. L. (1950). Development and appraisal of devices providing immediate automatic scoring of objective tests and concomitant self-instruction. *Journal of Psychology, 29,* 417–447.

Pressey, S. L. (1959). Certain major psycho-educational issues appearing in the conference on teaching machines. In E. Galanter (Ed.), *Automatic teaching: The state of the art* (pp. 187–198). New York: John Wiley & Sons.

Pressey, S. L. (1960). Some perspectives and major problems regarding teaching machines. In A. A. Lumsdaine & R. Glaser (Eds.), *Teaching machines and Programmed Instruction* (pp. 497–506). Washington, DC: National Education Association.

Price, J. E. (1963). Automated teaching programs with mentally retarded students. *American Journal of Mental Deficiency, 68,* 69–72.

Recommendations for reporting the effectiveness of Programmed Instructional materials. (1966). Washington, DC: National Education Association.

Reed, J. E., & Hayman, J. L. (1962). An experiment involving use of English 2600, an automated instruction text. *Journal of Educational Research, 55,* 476–484.

Reiser, R. A. (2001). A history of instructional design and technology: Part II: A history of instructional design. *Educational Technology Research and Development, 49*(2), 57–67.

Reiser, R. A., & Dempsey, J. A. (2002). *Trends and issues in instructional design and technology*. Upper Saddle River, NJ: Merrill/Prentice Hall.

Richmond, G. (1983). Comparison of automated and human instruction for developmentally retarded preschool children. *TASH Journal, 8,* 79–84.

Rigney, J. W., & Fry, E. B. (1961). Current teaching-machine programs and programming techniques. *Audiovisual Communication Review, 9*(3), Supplement 3, 7–121.

Roe, A., Massey, M., Weltman, G., & Leeds, D. (1960). *Automated teaching methods using linear programs*. UCLA Department of Engineering Report, 60–105.

Roe, A. A. (1960). *Automated teaching methods using linear programs* (Project number 60–105). Los Angeles, CA: University of California.

Roe, A. A. (1962). A comparison of branching methods for Programmed Instruction. *Journal of Educational Research, 55,* 407–416.

Roe, K. V., Case, H. W., & Roe, A. (1962). Scrambled vs. ordered sequence in auto-instructional programs. *Journal of Educational Psychology, 53,* 101–104.

Romiszowski, A. J. (1986). *Developing auto-instructional materials: From programmed texts to CAL and interactive video. Instructional Development 2*. London: Kogan Page.

Roth, R. H. (1963). Student reactions to programmed learning. *Phi Delta Kappan, 44,* 278–281.

Rothkokpf, E. Z. (1960). Some research problems in the design of materials and devices for automated teaching: In A. A. Lumsdaine & R. Glaser (Eds.), *Teaching machines and programmed learning* (pp. 318–328). Washington, DC: National Education Association.

Russo, D. C., Koegel, R. L., & Lovaas, O. I. (1978). A comparison of human and automated instruction of autistic children. *Journal of Abnormal Child Psychology, 6,* 189–201.

Rutkaus, M. A. (1987). Remember programmed instruction? *Educational Technology, 27*(10), 46–48.

Scharf, E. S. (1961). A study of the effects of partial reinforcement on behavior in a programmed learning situation. In R. Glaser & J. I. Taber, (Eds.), *Investigations of the characteristics of programmed learning sequences* (Research project number 691). Pittsburgh, PA: University of Pittsburgh.

Scriven, M. (1967). The methodology of evaluation. In *Perspectives of curriculum evaluation* (American Educational Research Association Monograph Series on Curriculum Evaluation, no. 1). Chicago: Rand McNally.

Severin, D. G. (1960). Appraisal of special tests and procedures used with self-scoring instructional testing devices. In A. A. Lumsdaine & R. Glaser (Eds.), *Teaching machines and programmed learning* (pp. 678–680). Washington, DC: National Education Association.

Shay, C. B. (1961). Relationship of intelligence to step size on a teaching machine program. *Journal of Educational Psychology, 52,* 93–103.

Shimamune, S. (1992). *Experimental and theoretical analysis of instructional tasks: reading, discrimination and construction.* Unpublished dissertation, W. Michigan University, Kalamazoo, MI.

Silberman, H. F. (1962). Characteristics of some recent studies of instructional methods. In J. E. Coulson (Ed.), *Programmed learning and computer-based instruction* (pp. 13–24), New York: John Wiley & Sons.

Silberman, H. F., Melaragno, R. J., Coulson, J. E., & Estavan, D. (1961). Fixed sequence versus branching auto-instructional methods. *Journal of Educational Psychology, 52,* 166–172.

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

Skinner, B. F. (1954). The science of learning and the art of teaching. *Harvard Educational Review, 24,* 86–97.

Skinner, B. F. (1957). *Verbal behavior*. New York: Appleton-Century Crofts.

Skinner, B. F. (1958a). Teaching machines. *Science, 128,* 969–977.

Skinner, B. F. (1958b). Reinforcement today. *American Psychologist, 13,* 94–99.

Skinner, B. F. (1961). Why we need teaching machines. *Harvard Educational Review, 31,* 377–398.

Skinner, B. F. (1963). Operant behavior. *American Psychologist, 18,* 503–515.

Skinner, B. F. (1968a). *The technology of teaching.* New York: Appleton-Century-Crofts Educational Division, Meredith Corporation.

Skinner, B. F. (1968b). Reflections on a decade of teaching machines. In R. A. Weisgerber (Ed.), *Instructional process and media innovation* (pp. 404–417). Chicago: Rand McNally & Co.

Skinner, B. F. (1986). Programmed instruction revisited. *Phi Delta Kappan, 68*(2),103–110.

Skinner, B. F., & Holland, J. G. (1960). The use of teaching machines in college instruction. In A. A. Lumsdaine & R. Glaser (Eds.), *Teaching machines and programmed learning* (pp. 159–172). Washington, DC: National Education Association.

Smith, K. U., & Smith, M. F. (1966). *Cybernetic principles of learning and educational design*. New York: Holt, Rinehart & Winston.

Smith, N. H. (1962). The teaching of elementary statistics by the conventional classroom method versus the method of Programmed Instruction. *Journal of Educational Research, 55,* 417–420.

Smith, P. L., & Ragan, T. J. (1999). *Instructional design (2nd ed.).* Upper Saddle River, NJ: Merrill Prentice Hall.

Smith, W., & Moore, J. W. (1962). Size-of-step and achievement in programmed spelling. *Psychological Reports, 10,* 287–294.

Snygg, D. (1962). The tortuous path of learning theory. *Audiovisual Instruction, 7,* 8–12.

Stephens, A. L. (1960). Certain special factors involved in the law of effect. In A. A. Lumsdaine & R. Glaser (Eds.), *Teaching machines and programmed learning* (pp. 89–93). Washington, DC: National Education Association.

Stewart, D., & Chown, S. (1965). A comparison of the effects of a continuous and a linear programmed text on adult pupils. *Occupational Psychology, 39,* 135.

Stickell, D. W. (1963). *A critical review of the methodology and results of research comparing televised and face-to-face instruction*. Unpublished doctoral dissertation, State College, PA: Pennsylvania State University.

Stolurow, L. M. (1961). *Teaching by machine*. Washington, DC: US Department of Health, Education & Welfare.

Stolurow, L. M. (1963). Programmed Instruction for the mentally retarded. *Review of Educational Research, 33,* 126–133.

Stolurow, L. M., Hasterok, S. G., & Ferrier, A. (1960). *Automation in education.* Unpublished paper presented at Allerton Park, IL Conference.

Stolurow, L. M., Peters, S., & Steinberg, M. (1960, October). *Prompting, confirmation, over learning, and retention.* Paper presented at the meeting of the Illinois Council of Exceptional Children, Chicago, IL.

Stolurow, L. M., & Walker, C. C. (1962). A comparison of overt and covert response in programmed learning. *Journal of Educational Research, 55,* 421–429.

Suppes, P., & Ginsburg, R. (1962). Experimental studies of mathematical concept formation in young children. *Science Education, 46,* 230–240.

Suppes, P., & Morningstar, M. (1969). Computer-assisted instruction. *Science, 166,* 343–350.

Taber, J. I., Glaser, R., & Schaefer, H. H. (1965). *Learning and Programmed Instruction.* Reading, MA: Addison-Wesley.

Tiemann, P. S., & Markle, S. M. (1990). *Analyzing instructional content: A guide to instruction and evaluation* (4th ed.) IL: Stipes Publishing Company.

Tillman, S., & Glynn, S. (1987). Writing text that teaches: Historical overview. *Educational Technology, 27*(10), 41–45.

Tobias, S. (1969a). Effect of attitudes to Programmed Instruction and other media on achievement from programmed materials. *AV Communication Review, 17*(13), 299–306.

Tobias, S. (1969b). Effect of creativity, response mode and subject matter familiarity on achievement from Programmed Instruction. *Journal of Educational Psychology, 60,* 453–460.

Tobias, S. (1973). Review of the response mode issue. *Review of Educational Research, 43,* 193–204.

Tobias, S., & Weiner, M. (1963). Effect of response made on immediate and delayed recall from programmed material. *Journal of Programed Instruction, 2,* 9–13.

Tsai, S. W., & Pohl, N. F. (1978). Student achievement in computer programming: Lecture vs. computer-aided instruction. *Journal of Experimental Education, 46,* 66–70.

Tudor, R. M. (1995). Insolating the effects of active responding in computer-based instruction. *Journal of Applied Behavior Analysis, 28,* 343–344.

Tudor, R. M. & Bostow, D. E. (1991). Computer-Programmed Instruction: The relation or required interaction to practical application. *Journal of Applied Behavior Analysis, 24,* 361–368.

Tyler, R. W. (1932). The construction of examinations in botany and zoology. *Service Studies in Higher Education* (pp. 49–50). Ohio State University Studies, Bureau of Educational Research Monographs, No. 15.

Univin, D. (1966). An organizational explanation for certain retention and correlation factors in a comparison between two teaching methods. *Programmed Learning and Educational Technology, 3,* 35–39.

Uttal, W. R. (1962). *My teacher has three arms!!* IBM Corporation, T. J. Watson Research paper, RC-788.

VanAtta, L. (1961). Behavior in small steps. *Contemporary Psychology, 6,* 378–381.

Vargas, E. A., & Vargas, J. (1992). Programmed instruction and teaching machines. In *Designs for excellence in education: The legacy of B. F. Skinner*. Longmont, CO: Sopris West, Inc.

Vunovick, P. (1995). *Discrimination training, terminal-response training and concept learning in the teaching of goal-directed-systems design.* Unpublished master's thesis, W. Michigan University, Kalamazoo, MI.

Wager, W. W., & Broaderick, W. A. (1974). Three objective rules of sequencing applied to programmed learning materials. *AV Communication Review, 22*(4), 423–438.

Weinstock, H. R., Shelton, F. W., & Pulley, J. L. (1973). Critique of criticism of CAI. *Educational Forum, 37,* 427–433.

Wendt, P. R., & Rust, G. (1962). Pictorial and performance frames in branching Programmed Instruction. *Journal of Educational Research, 55,* 430–432.

Widlake, P. (1964). English sentence construction: The effects of mode of response on learning. *Education Review, 16,* 120–129.

Williams, J. P. (1963). A comparison of several response modes in a review program. *Journal of Educational Psychology, 54,* 253–260.

Wittrock, M. C. (1963). Response mode in the programming of kinetic molecular theory concepts. *Journal of Educational Psychology, 54,* 89–93.

Wodkte, K. H., Brown, B. R., Sands, H. R., & Fredericks, P. (1968, February). *The effects of subject matter and individual difference variables on learning from scrambled versus ordered instructional programs.* Paper presented at the meeting of the American Educational Research Association, Chicago, IL.